1. Introduction
    1. Summary of Basic Rules for Probability Theory
    2. Decibel scale with signal processing applications
    3. Foundations of Probability Theory: Basic Definitions
    4. Review of Probability Theory
    5. Review of Linear Algebra
    6. Signal Power
    7. Signal Energy vs. Signal Power
    8. Introduction to Statistical Signal Processing
    9. Introduction to Random Signals and Processes
    10. Introduction to Stochastic Processes
    11. The Gaussian Process
    12. Sampling and Random Sequences
    13. Stationary and Nonstationary Random Processes
    14. Correlation and Covariance of a Random Signal
    15. Autocorrelation of Random Processes
    16. Crosscorrelation of Random Processes
    17. Cauchy-Schwarz Inequality
    18. The Q-function
2. Detection Theory
    1. Introduction to Detection Theory
    2. Discrete-Time Detection Theory
    3. Hypothesis Testing
    4. The Neyman-Pearson Criterion
    5. White Gaussian Noise
    6. Elementary Hypothesis Testing
    7. The Bayes Risk Criterion in Hypothesis Testing
    8. Sufficient Statistics
    9. Detection in the Presence of Unknowns
    10. Random Parameters
    11. Non-Random Parameters

Summary of Basic Rules for Probability Theory

“Probability theory is nothing but common sense reduced to calculation” (Laplace).

Introduction

This module was adapted from E.T. Jaynes’ manuscript entitled: “Probability Theory with Applications to Science and Engineering – A Series of Informal Lectures”, 1974. The entire manuscript is available at http://bayes.wustl.edu/etj/science.pdf.html.

A second and significantly expanded edition of this manuscript is available on Amazon. The first 3 chapters of the second edition are available here http://bayes.wustl.edu/etj/prob/book.pdf.

Deductive Logic (Boolean Algebra)

Denote propositions by A, B, etc., their denials by A_c, B_c etc. Define the logical product and logical sum by

$\mathrm{AB} \equiv$ “Both A and B are true”

$A + B \equiv$ “At least one of the propositions, A, B are true”

Deductive reasoning then consists of applying relations such as

$A + A = A$;

$\mathrm{A}(B + C) = (\mathrm{AB}) + (\mathrm{AC})$;

if $\mathrm{D} = \mathrm{A}_c B_\mathrm{c}$then $D_c = \mathrm{A} + B$.

Inductive Logic (Probability Theory)

Inductive logic is the extension of deductive logic, describing the reasoning of an idealized “robot”, who represents degrees of plausibility of a logical proposition by real numbers:

$p(A \mid B)$= probability of A, given B.

We use the original term “robot” advocated by Jaynes, it is intended to mean the use of inductive logic that follows a set of consistent rules that can be agreed upon. In this formulation of probability theory, conditional probabilities are fundamental. The elementary requirements of common sense and consistency determine these basic rules of reasoning (see Jaynes for the derivation).

In these rules, one can think of the proposition C being the prior information that is available to assign probabilities to logical propositions, but these rules are true without this interpretation.

Rule 1: $p(\mathrm{AB} \mid C) = \mathrm{p}(A \mid \mathrm{BC})p(B \mid C) = p(B \mid \mathrm{AC})p(A \mid C)$

Rule 2: $p(A \mid B) + p(A_c \mid B) = 1$

Rule 3: $p(A + B \mid C) = p(A \mid C) + p(B \mid C) - p(\mathrm{AB} \mid C)$

Rule 4: If $\{A_1, \ldots A_N\}$are mutually exclusive and exhaustive, and information $B$is indifferent to tem; i.e. if B gives no preference to one over any other then:

$p(A_i \mid B) = 1/n$,$i = 1 \ldots n$ (principle of insufficient reason)

From rule 1 we obtain Bayes’ theorem:

$p(A \mid \mathrm{BC}) = p(A \mid C)\frac{p(B|\mathrm{AC})}{p(B|C)}$

From Rule 3, if $\{A_1, \ldots A_N\}$are mutually exclusive,

$p(A_1 + \ldots A_N \mid B) = \sum_{i=1}^{n} p(A_i \mid B)$

If in addition, the $A_i$are exhaustive, we obtain the chain rule:

$p(B \mid C) = \sum_{i=1}^{n} p(\mathrm{BA}_i \mid C) = \sum_{i=1}^{n} p(B \mid A_i C) p(A_i \mid C)$

Prior Probabilities

The initial information available to the robot at the beginning of any problem is denoted by X. $p(A \mid X)$is then the prior probability of A. Applying Bayes' theorem to take account of new evidence Eyields the posterior probability $p(A \mid \mathrm{EX})$. In a posterior probability we sometimes leave off the X for brevity: $p(A \mid E) \equiv p(A \mid \mathrm{EX})$.

Prior probabilities are determined by Rule 4 when applicable; or more generally by the principle of maximum entropy.

Decision Theory

Enumerate the possible decisions $D_1, \ldots D_k$and introduce the loss function $L(D_i, \theta_i)$representing the "loss" incurred by making decision $D_i$if $\theta_j$is the true state of nature. After accumulating new evidence E, make that decision $D_i$which minimizes the expected loss over the posterior distribution of θ_j :

Choose the decision $D_i$which minimizes $\langle L \rangle_i = \sum_j L(D_i, \theta_j) p(\theta_j \mid \mathrm{EX})$

$\mathrm{choose} D_i \mathrm{such\ that\ is\ minimized}$

Decibel scale with signal processing applications

Introduction

The concept of decibel originates from telephone engineers who were working with power loss in a telephone line consisting of cascaded circuits. The power loss in each circuit is the ratio of the power in to the power out, or equivivalently, the power gain is the ratio of the power out to the power in.

Let P_{in} be the power input to a telephone line and P_{out} the power out. The power gain is then given by
Equation:

$$\text{Gain} = \frac{P_{\text{out}}}{P_{\text{in}}}$$

Taking the logarithm of the gain formula we obtain a comparative measure called Bel.

> **Note:** $\text{Gain (Bel)} = \log \frac{P_{\text{out}}}{P_{\text{in}}}$

This measure is in honour of Alexander G. Bell, see [link].

Alexander G. Bell

Decibel

Bel is often a to large quantity, so we define a more useful measure, decibel:
Equation:

$$\text{Gain (dB)} = 10 \log \frac{P_{\text{out}}}{P_{\text{in}}}$$

Please note from the definition that the gain in dB is relative to the input power. In general we define:
Equation:

$$\text{Number of decibels} = 10 \log \frac{P}{P_{\text{ref}}}$$

If no reference level is given it is customary to use $P_{\text{ref}} = 1 \text{ W}$, in which case we have:

Note: $\text{Number of decibels} = 10 \log P$

Example:
Given the power spectrum density (psd) function of a signal $x(n)$, $S_{\text{xx}}(if)$.
Express the magnitude of the psd in decibels.
We find $S_{\text{xx}}(\text{dB}) = 10 \log |S_{\text{xx}}(if)|$.

More about decibels

Above we've calculated the decibel equivalent of power. Power is a quadratic variable, whereas voltage and current are linear variables. This can be seen, for example, from the formulas $P = \frac{V^2}{R}$ and $P = I^2 R$.

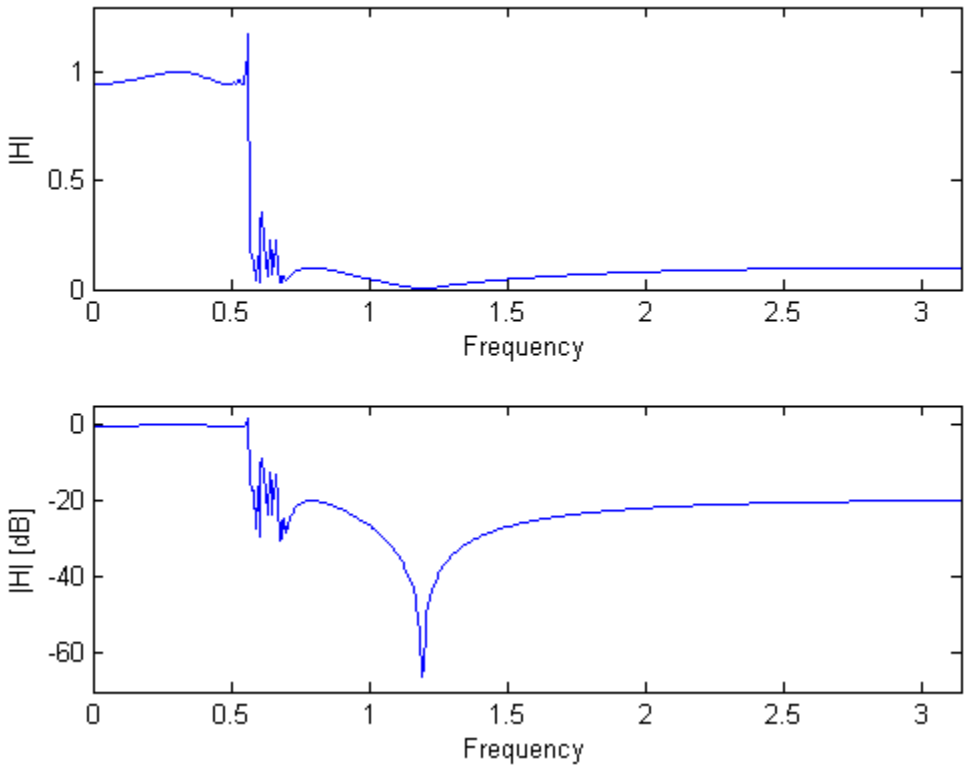
So if we want to find the decibel value of a current or voltage, or more general an amplitude we use:
Equation:

$$\text{Amplitude}\,(\text{dB}) = 20 \log \frac{\text{Amplitude}}{\text{Amplitude}_{\text{ref}}}$$

This is illustrated in the following example.

Example:
Express the magnitude of the filter $H(if)$ in dB scale.
The magnitude is given by $|H(if)|$, which gives:
$|H(\text{dB})| = 20 \log |H(if)|$.

Plots of the magnitude of an example filter $|H(if)|$ and its decibel equivalent are shown in [link].



Magnitude responses.

Some basic arithmetic

The ratios 1,10,100, 1000 give dB values 0 dB, 10 dB, 20 dB and 30 dB respectively. This implies that an increase of 10 dB corresponds to a ratio increase by a factor 10.

This can easily be shown: Given a ratio R we have R[dB] = 10 log R. Increasing the ratio by a factor of 10 we have: 10 log (10*R) = 10 log 10 + 10 log R = 10 dB + R dB.

Another important dB-value is 3dB. This comes from the fact that:

An increase by a factor 2 gives: an increase of 10 log 2 ≈ 3 dB. A “increase” by a factor 1/2 gives: an “increase” of 10 log 1/2 ≈ -3 dB.

Example:
In filter terminology the **cut-off frequency** is a term that often appears. The cutoff frequency (for lowpass and highpass filters), f_c, is the frequency at which the squared magnitude response in dB is ½. In decibel scale this corresponds to about -3 dB.

Decibels in linear systems

In signal processing we have the following relations for linear systems:
Equation:

$$Y(if) = H(if)X(if)$$

where X and H denotes the input signal and the filter respectively. Taking absolute values on both sides of [link] and converting to decibels we get:

Note: The output amplitude at a given frequency is simply given by the sum of the filter gain and the input amplitude, both in dB.

Other references:

Above we have used $P_{\text{ref}} = 1$ W as a reference and obtained the standard dB measure. In some applications it is more useful to use $P_{\text{ref}} = 1$ mW and we then have the dBm measure.

Another example is when calculating the gain of different antennas. Then it is customary to use an isotropic (equal radiation in all directions) antenna as a reference. So for a given antenna we can use the dBi measure. (i -> isotropic)

Matlab files

filter_example.m

Foundations of Probability Theory: Basic Definitions

Basic Definitions

The basis of probability theory is a set of events - sample space - and a systematic set of numbers - probabilities - assigned to each event. The key aspect of the theory is the system of assigning probabilities. Formally, a **sample space** is the set Ω of all possible outcomes ω_i of an experiment. An **event** is a collection of sample points ω_i determined by some set-algebraic rules governed by the laws of Boolean algebra. Letting A and B denote events, these laws are

$$\text{Union:} A \cup B = \{\omega | \, \omega \in A \ \vee \ \omega \in B\}$$

$$\text{Intersection:} A \cap B = \{\omega | \, \omega \in A \ \wedge \ \omega \in B\}$$

$$\text{Complement:} A' = \{\omega | \, \omega \notin A\}$$

$$(A \cup B)' = A' \cap B'$$

The null set $\emptyset$ is the complement of Ω. Events are said to be **mutually exclusive** if there is no element common to both events: $A \cap B = \emptyset$.

Associated with each event A_i is a **probability measure** $\Pr[A_i]$, sometimes denoted by π_i, that obeys the **axioms of probability**.

- $\Pr[A_i] \geq 0$
- $\Pr[\Omega] = 1$
- If $A \cap B = \emptyset$, then $\Pr[A \cup B] = \Pr[A] + \Pr[B]$.

The consistent set of probabilities $\Pr[\cdot]$ assigned to events are known as the **a priori probabilities**. From the axioms, probability assignments for Boolean expressions can be computed. For example, simple Boolean manipulations ($A \cup B = A \cup A'B$) lead to
Equation:

$$\Pr[A \cup B] = \Pr[A] + \Pr[B] - \Pr[A \cap B]$$

Suppose $\Pr[B] \neq 0$. Suppose we know that the event B has occurred; what is the probability that event A has also occurred? This calculation is known as the **conditional probability** of A given B and is denoted by $\Pr[A \mid B]$. To evaluate conditional probabilities, consider B to be the sample space rather than Ω. To obtain a probability assignment under these circumstances consistent with the axioms of probability, we must have
Equation:

$$\Pr[A \mid B] = \frac{\Pr[A \cap B]}{\Pr[B]}$$

The event is said to be **statistically independent** of B if
$\Pr[A \mid B] = \Pr[A]$: the occurrence of the event B does not change the probability that A occurred. When independent, the probability of their intersection $\Pr[A \cap B]$ is given by the product of the a priori probabilities $\Pr[A] \Pr[B]$. This property is necessary and sufficient for the independence of the two events. As $\Pr[A \mid B] = \frac{\Pr[A \cap B]}{\Pr[B]}$ and $\Pr[B \mid A] = \frac{\Pr[A \cap B]}{\Pr[A]}$, we obtain **Bayes' Rule**.
Equation:

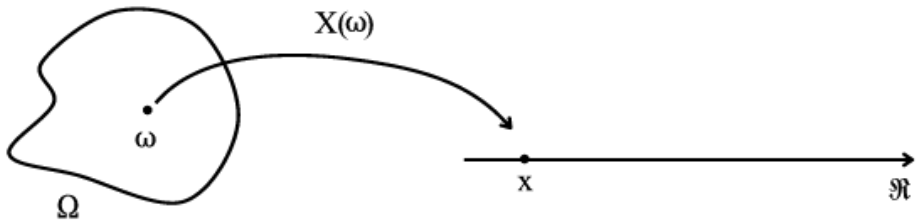
$$\Pr[B \mid A] = \frac{\Pr[A \mid B] \Pr[B]}{\Pr[A]}$$

Review of Probability Theory

The focus of this course is on digital communication, which involves transmission of information, in its most general sense, from source to destination using digital technology. Engineering such a system requires modeling both the information and the transmission media. Interestingly, modeling both digital or analog information and many physical media requires a probabilistic setting. In this chapter and in the next one we will review the theory of probability, model random signals, and characterize their behavior as they traverse through deterministic systems disturbed by noise and interference. In order to develop practical models for random phenomena we start with carrying out a random experiment. We then introduce definitions, rules, and axioms for modeling within the context of the experiment. The outcome of a random experiment is denoted by ω. The sample space Ω is the set of all possible outcomes of a random experiment. Such outcomes could be an abstract description in words. A scientific experiment should indeed be repeatable where each outcome could naturally have an associated probability of occurrence. This is defined formally as the ratio of the number of times the outcome occurs to the total number of times the experiment is repeated.

Random Variables

A random variable is the assignment of a real number to each outcome of a random experiment.



Example:
Roll a dice. Outcomes $\{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_6\}$
ω_i = i dots on the face of the dice.
$X(\omega_i) = i$

Distributions

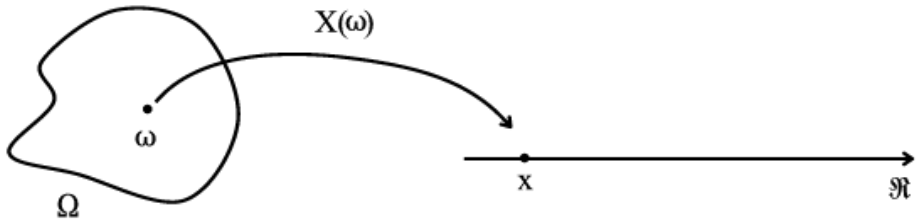
Probability assignments on intervals $a < X \leq b$

Cumulative distribution

The cumulative distribution function of a random variable X is a function $\mathrm{F}_X\left(\mathbb{R} \mapsto \mathbb{R}\right)$ such that

Equation:

$$\begin{aligned} \mathrm{F}_X\left(b\right) &= \Pr[X \leq b] \\ &= \Pr[\{\omega \in \Omega | \, X(\omega) \leq b\}] \end{aligned}$$



Continuous Random Variable

A random variable X is continuous if the cumulative distribution function can be written in an integral form, or

Equation:

$$\mathrm{F}_X\left(b\right) = \int_{-\infty}^{b} \mathrm{f}_X\left(x\right) \mathrm{d}\, x$$

and $\mathrm{f}_X\left(x\right)$ is the probability density function (pdf) (e.g., $\mathrm{F}_X\left(x\right)$ is differentiable and $\mathrm{f}_X\left(x\right) = \frac{\mathrm{d}}{\mathrm{d}x}(\mathrm{F}_X\left(x\right))$)

Discrete Random Variable

A random variable X is discrete if it only takes at most countably many points (i.e., $\mathrm{F}_X\left(\cdot\right)$ is piecewise constant). The probability mass function (pmf) is defined as

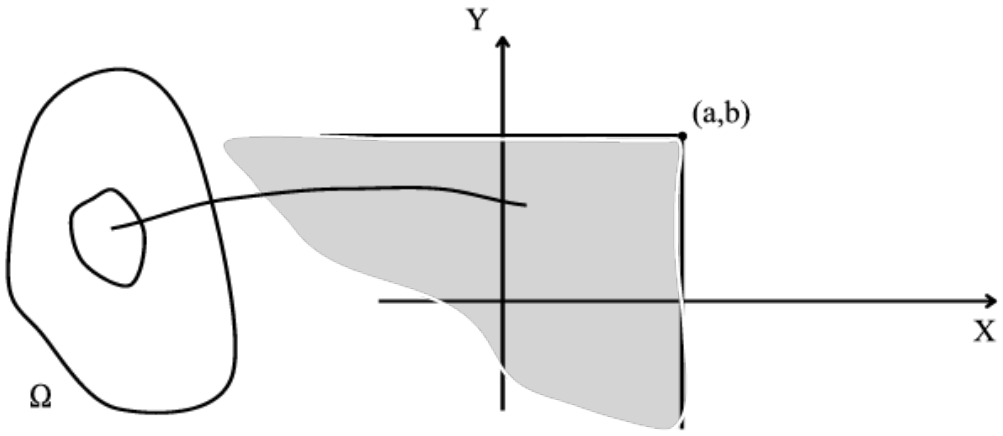
Equation:

$$\begin{aligned} \mathrm{p}_X\left(x_k\right) &= \Pr[X = x_k] \\ &= \mathrm{F}_X\left(x_k\right) - \operatorname*{limit}_{x(x \to x_k) \wedge (x < x_k)} \mathrm{F}_X\left(x\right) \end{aligned}$$

Two random variables defined on an experiment have joint distribution

Equation:

$$\begin{aligned} \mathrm{F}_{X,,Y}\left(a,b\right) &= \Pr[X \leq a, Y \leq b] \\ &= \Pr[\{\omega \in \Omega |\ \left(X(\omega) \leq a\right)\ \wedge\ \left(Y(\omega) \leq b\right)\}] \end{aligned}$$



Joint pdf can be obtained if they are jointly continuous
Equation:

$$\mathrm{F}_{X,,Y}\left(a,b\right) = \int_{-\infty}^{b}\int_{-\infty}^{a}\mathrm{f}_{X,Y}\left(x,y\right)\,\mathrm{d}\,x\,\mathrm{d}\,y$$

(e.g., $\mathrm{f}_{X,Y}\left(x,y\right) = \frac{\partial^2 \mathrm{F}_{X,,Y}(x,y)}{\partial x \partial y}$)

Joint pmf if they are jointly discrete
Equation:

$$\mathrm{p}_{X,Y}\left(x_k,y_l\right) = \Pr[X = x_k, Y = y_l]$$

Conditional density function
Equation:

$$f_{Y|X}(y|x) = \frac{\mathrm{f}_{X,Y}\left(x,y\right)}{\mathrm{f}_X\left(x\right)}$$

for all x with $\mathrm{f}_X\ (x) > 0$ otherwise conditional density is not defined for those values of x with $\mathrm{f}_X\ (x) = 0$

Two random variables are **independent** if
Equation:

$$\mathrm{f}_{X,Y}\ (x, y) = \mathrm{f}_X\ (x)\ \mathrm{f}_Y\ (y)$$

for all $x \in \mathbb{R}$ and $y \in \mathbb{R}$. For discrete random variables,
Equation:

$$\mathrm{p}_{X,Y}\ (x_k, y_l) = \mathrm{p}_X\ (x_k)\ \mathrm{p}_Y\ (y_l)$$

for all k and l.

Moments

Statistical quantities to represent some of the characteristics of a random variable.
Equation:

$$\begin{aligned} \overline{g(X)} &= E[g(X)] \\ &= \begin{cases} \int_{-\infty}^{\infty} g(x)\ \mathrm{f}_X\ (x)\ \mathrm{d}\ x & \text{if continuous} \\ \sum_k g(x_k)\ \mathrm{p}_X\ (x_k) & \text{if discrete} \end{cases} \end{aligned}$$

- Mean
  Equation:

$$\mu_X = \overline{X}$$

- Second moment
  Equation:

$$E\left[X^2\right] = \overline{X^2}$$

- Variance
  Equation:

$$\begin{aligned}\operatorname{Var}(X) &= \sigma(X)^2 \\ &= \overline{(X-\mu_X)^2} \\ &= \overline{X^2} - {\mu_X}^2\end{aligned}$$

- Characteristic function
  Equation:

$$\Phi_X(u) = \overline{e^{iuX}}$$

  for $u \in \mathbb{R}$, where $i = \sqrt{-1}$
- Correlation between two random variables
  Equation:

$$\begin{aligned}R_{XY} &= \overline{XY^*} \\ &= \begin{cases}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy^*\ \mathrm{f}_{X,Y}\left(x,y\right)\,\mathrm{d}\,x\,\mathrm{d}\,y & \text{if X and Y are jointly continuous} \\ \sum_k\sum_l x_k y_l^*\ \mathrm{p}_{X,Y}\left(x_k,y_l\right) & \text{if X and Y are jointly discrete}\end{cases}\end{aligned}$$

- Covariance
  Equation:

$$\begin{aligned}C_{XY} &= \operatorname{Cov}(X,Y) \\ &= \overline{(X-\mu_X)(Y-\mu_Y)^*} \\ &= R_{XY} - \mu_X\mu_Y^*\end{aligned}$$

- Correlation coefficient
  Equation:

$$\rho_{XY} = \frac{\operatorname{Cov}(X,Y)}{\sigma_X\sigma_Y}$$

Uncorrelated random variables

Two random variables X and Y are uncorrelated if $\rho_{XY} = 0$.

Review of Linear Algebra

Vector spaces are the principal object of study in linear algebra. A vector space is always defined with respect to a field of scalars.

Fields

A field is a set F equipped with two operations, addition and mulitplication, and containing two special members 0 and 1 ($0 \neq 1$), such that for all $\{a, b, c\} \in F$

1. 1. $(a + b) \in F$
   2. $a + b = b + a$
   3. $(a + b) + c = a + (b + c)$
   4. $a + 0 = a$
   5. there exists $-a$ such that $a + -a = 0$

2. 1. $ab \in F$
   2. $ab = ba$
   3. $(ab)c = a\,(bc)$
   4. $a \cdot 1 = a$
   5. there exists a^{-1} such that $aa^{-1} = 1$

3. $a\,(b + c) = ab + ac$

More concisely

1. F is an abelian group under addition
2. F is an abelian group under multiplication
3. multiplication distributes over addition

Examples

$\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$

Vector Spaces

Let F be a field, and V a set. We say V **is a vector space over** F if there exist two operations, defined for all $a \in F$, $\boldsymbol{u} \in V$ and $\boldsymbol{v} \in V$:

- vector addition: $(\boldsymbol{u}, \boldsymbol{v}) \to (\boldsymbol{u} + \boldsymbol{v}) \in V$
- scalar multiplication: $(a, \boldsymbol{v}) \to a\boldsymbol{v} \in V$

and if there exists an element denoted $\mathbf{0} \in V$, such that the following hold for all $a \in F$, $b \in F$, and $\boldsymbol{u} \in V$, $\boldsymbol{v} \in V$, and $\boldsymbol{w} \in V$

1. 1. $\boldsymbol{u} + (\boldsymbol{v} + \boldsymbol{w}) = (\boldsymbol{u} + \boldsymbol{v}) + \boldsymbol{w}$
   2. $\boldsymbol{u} + \boldsymbol{v} = \boldsymbol{v} + \boldsymbol{u}$
   3. $\boldsymbol{u} + \mathbf{0} = \boldsymbol{u}$
   4. there exists $-\boldsymbol{u}$ such that $\boldsymbol{u} + -\boldsymbol{u} = \mathbf{0}$

2. 1. $a\,(\boldsymbol{u} + \boldsymbol{v}) = a\boldsymbol{u} + a\boldsymbol{v}$
   2. $(a + b)\boldsymbol{u} = a\boldsymbol{u} + b\boldsymbol{u}$
   3. $(ab)\boldsymbol{u} = a\,(b\boldsymbol{u})$
   4. $1 \cdot \boldsymbol{u} = \boldsymbol{u}$

More concisely,

1. V is an abelian group under plus
2. Natural properties of scalar multiplication

Examples

- $\mathbb{R}^N$ is a vector space over $\mathbb{R}$
- $\mathbb{C}^N$ is a vector space over $\mathbb{C}$
- $\mathbb{C}^N$ is a vector space over $\mathbb{R}$
- $\mathbb{R}^N$ is **not** a vector space over $\mathbb{C}$

The elements of V are called **vectors**.

Euclidean Space

Throughout this course we will think of a signal as a vector

$$\boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{pmatrix} = \begin{pmatrix} x_1 & x_2 & \dots & x_N \end{pmatrix}^T$$

The samples $\{x_i\}$ could be samples from a finite duration, continuous time signal, for example.

A signal will belong to one of two vector spaces:

Real Euclidean space

$\boldsymbol{x} \in \mathbb{R}^N$ (over $\mathbb{R}$)

Complex Euclidean space

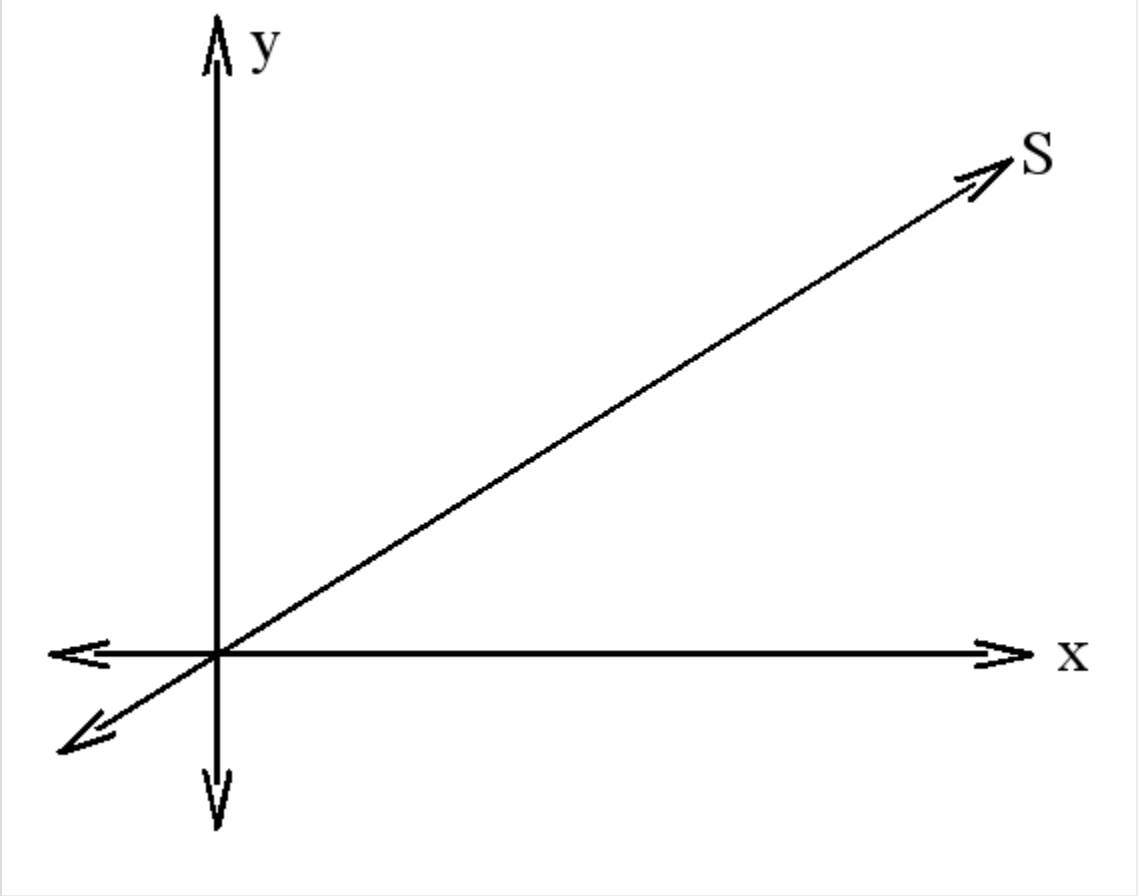
$\boldsymbol{x} \in \mathbb{C}^N$ (over $\mathbb{C}$)

Subspaces

Let V be a vector space over F.

A subset $S \subseteq V$ is called a **subspace** of V if S is a vector space over F in its own right.

Example:
$V = \mathbb{R}^2$, $F = \mathbb{R}$, $S =$ any line though the origin.



S is any line through the origin.

Are there other subspaces?

$S \subseteq V$ is a subspace if and only if for all $a \in F$ and $b \in F$ and for all $\boldsymbol{s} \in S$ and $\boldsymbol{t} \in S$, $(a\boldsymbol{s} + b\boldsymbol{t}) \in S$

Linear Independence

Let $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k \in V$.

We say that these vectors are **linearly dependent** if there exist scalars $a_1, \ldots, a_k \in F$ such that

Equation:

$$\sum_{i=1}^{k} a_i u_i = \mathbf{0}$$

and at least one $a_i \neq 0$.

If [link] only holds for the case $a_1 = \ldots = a_k = 0$, we say that the vectors are **linearly independent**.

Example:

$$1\begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix} - 2\begin{pmatrix} -2 \\ 3 \\ 0 \end{pmatrix} + 1\begin{pmatrix} -5 \\ 7 \\ -2 \end{pmatrix} = \mathbf{0}$$

so these vectors are linearly dependent in $\mathbb{R}^3$.

Spanning Sets

Consider the subset $S = \{v_1, v_2, \ldots, v_k\}$. Define the **span** of S

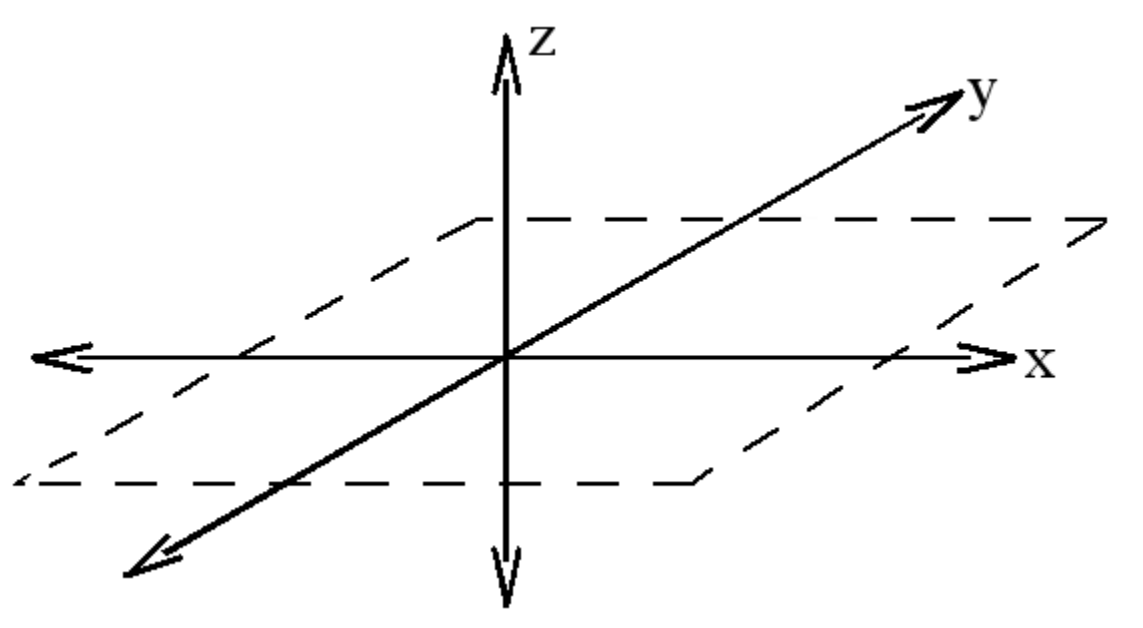
$$< S > \equiv \text{span}\,(S) \equiv \left\{ \sum_{i=1}^{k} a_i v_i \,\middle|\, a_i \in F \right\}$$

Fact: $< S >$ is a subspace of V.

Example:

$V = \mathbb{R}^3$, $F = \mathbb{R}$, $S = \{v_1, v_2\}$, $v_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$, $v_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \Rightarrow$

$< S >=$ xy-plane.



$< S >$ is the xy-plane.

Aside

If S is infinite, the notions of linear independence and span are easily generalized:

We say S is linearly independent if, for every finite collection $u_1, \ldots, u_k \in S$, (k arbitrary) we have

$$\left(\sum_{i=1}^{k} a_i u_i = \mathbf{0} \right) \Rightarrow \forall i : (a_i = 0)$$

The span of S is

$$< S >= \left\{ \sum_{i=1}^{k} a_i u_i \,\middle|\, a_i \in F \ \wedge \ u_i \in S \ \wedge \ (k < \infty) \right\}$$

Note: In both definitions, we only consider **finite** sums.

Bases

A set $B \subseteq V$ is called a **basis** for V over F if and only if

1. B is linearly independent
2. $< B >= V$

Bases are of fundamental importance in signal processing. They allow us to decompose a signal into building blocks (basis vectors) that are often more easily understood.

Example:
V = (real or complex) Euclidean space, $\mathbb{R}^N$ or $\mathbb{C}^N$.

$$B = \{e_1, \ldots, e_N\} \equiv \text{standard basis}$$

$$e_i = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix}$$

where the 1 is in the i^{th} position.

Example:
$V = \mathbb{C}^N$ over $\mathbb{C}$.

$$B = \{u_1, \ldots, u_N\}$$

which is the DFT basis.

$$u_k = \begin{pmatrix} 1 \\ e^{-\left(i2\pi\frac{k}{N}\right)} \\ \vdots \\ e^{-\left(i2\pi\frac{k}{N}(N-1)\right)} \end{pmatrix}$$

where $i = \sqrt{-1}$.

Key Fact

If B is a basis for V, then every $\boldsymbol{v} \in V$ can be written uniquely (up to order of terms) in the form

$$\boldsymbol{v} = \sum_{i=1}^{N} a_i v_i$$

where $a_i \in F$ and $v_i \in B$.

Other Facts

- If S is a linearly independent set, then S can be extended to a basis.
- If $< S >= V$, then S contains a basis.

Dimension

Let V be a vector space with basis B. The dimension of V, denoted $\dim(V)$, is the cardinality of B.

Every vector space has a basis.

Every basis for a vector space has the same cardinality.

$\Rightarrow \dim(V)$ is **well-defined**.

If $\dim(V) < \infty$, we say V is **finite dimensional**.

Examples

vector space	field of scalars	dimension
$\mathbb{R}^N$	$\mathbb{R}$	
$\mathbb{C}^N$	$\mathbb{C}$	
$\mathbb{C}^N$	$\mathbb{R}$	

Every subspace is a vector space, and therefore has its own dimension.

Example:
Suppose $(S = \{u_1, \ldots, u_k\}) \subseteq V$ is a linearly independent set. Then

$$\dim(< S >) =$$

Facts

- If S is a subspace of V, then $\dim(S) \leq \dim(V)$.
- If $\dim(S) = \dim(V) < \infty$, then $S = V$.

Direct Sums

Let V be a vector space, and let $S \subseteq V$ and $T \subseteq V$ be subspaces.

We say V is the **direct sum** of S and T, written $V = S \oplus T$, if and only if for every $\boldsymbol{v} \in V$, there exist unique $\boldsymbol{s} \in S$ and $\boldsymbol{t} \in T$ such that $\boldsymbol{v} = \boldsymbol{s} + \boldsymbol{t}$.

If $V = S \oplus T$, then T is called a **complement** of S.

Example:

$$V = C' = \{f : \mathbb{R} \to \mathbb{R} | f \text{is continuous}\}$$

$$S = \text{even funcitons in} C'$$

$$T = \text{odd funcitons in} C'$$

$$f(t) = \frac{1}{2}\left(f(t) + f(-t)\right) + \frac{1}{2}\left(f(t) - f(-t)\right)$$

If $f = g + h = g' + h'$, $g \in S$ and $g' \in S$, $h \in T$ and $h' \in T$, then $g - g' = h' - h$ is odd and even, which implies $g = g'$ and $h = h'$.

Facts

1. Every subspace has a complement
2. $V = S \oplus T$ if and only if
    1. $S \cap T = \{\mathbf{0}\}$
    2. $< S, T >= V$

3. If $V = S \oplus T$, and $\dim(V) < \infty$, then
   $\dim(V) = \dim(S) + \dim(T)$

Proofs

Invoke a basis.

Norms

Let V be a vector space over F. A norm is a mapping $V \to F$, denoted by $\| \cdot \|$, such that forall $\boldsymbol{u} \in V$, $\boldsymbol{v} \in V$, and $\lambda \in F$

1. $\| \boldsymbol{u} \| > 0$ if $\boldsymbol{u} \neq \boldsymbol{0}$
2. $\| \lambda \boldsymbol{u} \| = |\lambda| \, \| \boldsymbol{u} \|$
3. $\| \boldsymbol{u} + \boldsymbol{v} \| \leq \| \boldsymbol{u} \| + \| \boldsymbol{v} \|$

Examples

Euclidean norms:

$\boldsymbol{x} \in \mathbb{R}^N$:

$$\| \boldsymbol{x} \| = \left(\sum_{i=1}^{N} {x_i}^2 \right)^{\frac{1}{2}}$$

$\boldsymbol{x} \in \mathbb{C}^N$:

$$\| \boldsymbol{x} \| = \left(\sum_{i=1}^{N} \left(|x_i| \right)^2 \right)^{\frac{1}{2}}$$

Induced Metric

Every norm induces a metric on V

$$d(\boldsymbol{u}, \boldsymbol{v}) \equiv \| \boldsymbol{u} - \boldsymbol{v} \|$$

which leads to a notion of "distance" between vectors.

Inner products

Let V be a vector space over F, $F = \mathbb{R}$ or $\mathbb{C}$. An inner product is a mapping $V \times V \to F$, denoted $\langle \cdot, \cdot \rangle$, such that

1. $\langle \boldsymbol{v}, \boldsymbol{v} \rangle \geq 0$, and $\langle \boldsymbol{v}, \boldsymbol{v} \rangle = 0 \Leftrightarrow \boldsymbol{v} = 0$
2. $\langle \boldsymbol{u}, \boldsymbol{v} \rangle = \overline{\langle \boldsymbol{v}, \boldsymbol{u} \rangle}$
3. $\langle a\boldsymbol{u} + b\boldsymbol{v}, \boldsymbol{w} \rangle = a \langle (\boldsymbol{u}, \boldsymbol{w}) \rangle + b \langle (\boldsymbol{v}, \boldsymbol{w}) \rangle$

Examples

$\mathbb{R}^N$ over $\mathbb{R}$:

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \left(\boldsymbol{x}^T \boldsymbol{y} \right) = \sum_{i=1}^{N} x_i y_i$$

$\mathbb{C}^N$ over $\mathbb{C}$:

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \left(\boldsymbol{x}^{\mathrm{H}} \boldsymbol{y} \right) = \sum_{i=1}^{N} \overline{x_i} y_i$$

If $\left(\boldsymbol{x} = (x_1 \ldots x_N)^T \right) \in \mathbb{C}$, then

$$\boldsymbol{x}^{\mathrm{H}} \equiv \begin{pmatrix} \overline{x_1} \\ \vdots \\ \overline{x_N} \end{pmatrix}^T$$

is called the "Hermitian," or "conjugate transpose" of $\boldsymbol{x}$.

Triangle Inequality

If we define $\| \boldsymbol{u} \| = \langle \boldsymbol{u}, \boldsymbol{u} \rangle$, then

$$\| \boldsymbol{u} + \boldsymbol{v} \| \leq \| \boldsymbol{u} \| + \| \boldsymbol{v} \|$$

Hence, every inner product induces a norm.

Cauchy-Schwarz Inequality

For all $\boldsymbol{u} \in V$, $\boldsymbol{v} \in V$,

$$|\langle \boldsymbol{u}, \boldsymbol{v} \rangle| \leq \| \boldsymbol{u} \| \ \ \| \boldsymbol{v} \|$$

In inner product spaces, we have a notion of the angle between two vectors:

$$\left(\angle(\boldsymbol{u}, \boldsymbol{v}) = \arccos\left(\frac{\langle \boldsymbol{u}, \boldsymbol{v} \rangle}{\| \boldsymbol{u} \| \ \ \| \boldsymbol{v} \|} \right) \right) \in [0, 2\pi)$$

Orthogonality

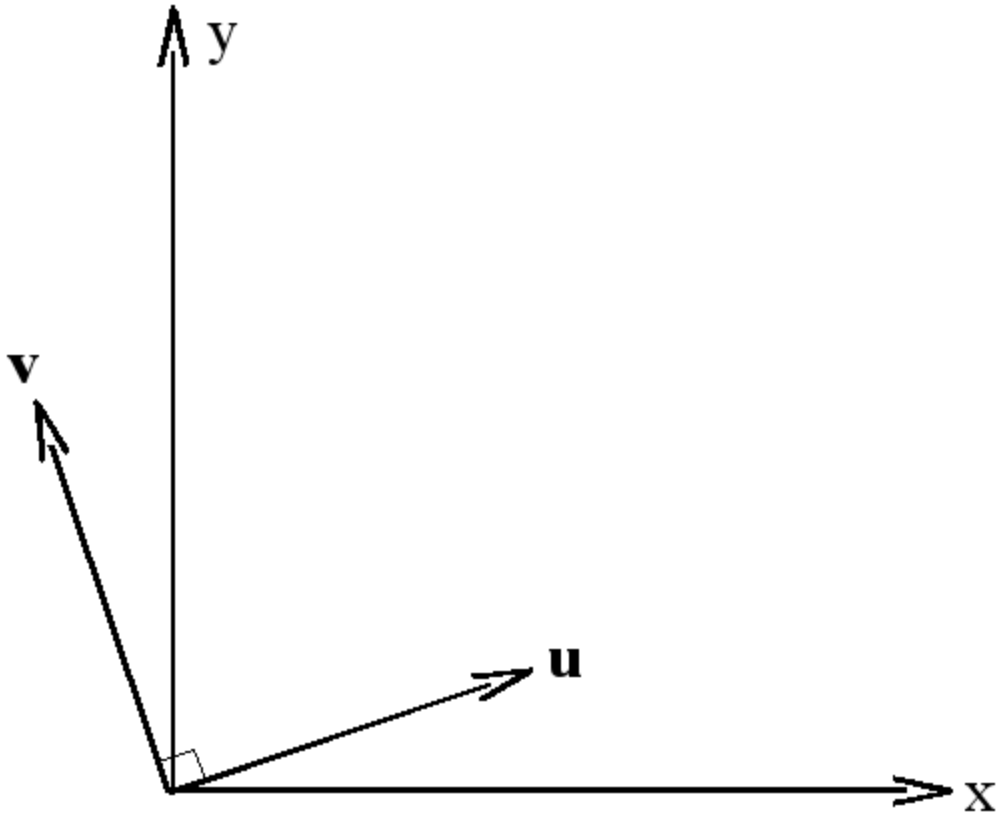
$\boldsymbol{u}$ and $\boldsymbol{v}$ are **orthogonal** if

$$\langle \boldsymbol{u}, \boldsymbol{v} \rangle = 0$$

Notation: $\boldsymbol{u} \perp \boldsymbol{v}$.

If in addition $\| \boldsymbol{u} \| = \| \boldsymbol{v} \| = 1$, we say $\boldsymbol{u}$ and $\boldsymbol{v}$ are **orthonormal**.

In an orthogonal (orthonormal) **set**, each pair of vectors is orthogonal (orthonormal).



Orthogonal vectors in $\mathbb{R}^2$.

Orthonormal Bases

An Orthonormal basis is a basis $\{v_i\}$ such that

$$\langle v_i, v_i \rangle = \delta_{ij} = \begin{cases} 1 & \text{if } \ i = j \\ 0 & \text{if } \ i \neq j \end{cases}$$

Example:
The standard basis for $\mathbb{R}^N$ or $\mathbb{C}^N$

Example:
The normalized DFT basis

$$u_k = \frac{1}{\sqrt{N}} \begin{pmatrix} 1 \\ e^{-\left(i2\pi \frac{k}{N}\right)} \\ \vdots \\ e^{-\left(i2\pi \frac{k}{N}(N-1)\right)} \end{pmatrix}$$

Expansion Coefficients

If the representation of $\boldsymbol{v}$ with respect to $\{v_i\}$ is

$$\boldsymbol{v} = \sum_i a_i v_i$$

then

$$a_i = \langle v_i, \boldsymbol{v} \rangle$$

Gram-Schmidt

Every inner product space has an orthonormal basis. Any (countable) basis can be made orthogonal by the Gram-Schmidt orthogonalization process.

Orthogonal Compliments

Let $S \subseteq V$ be a subspace. The **orthogonal compliment** S is

$$S^{\perp} = \{\boldsymbol{u} | \, \boldsymbol{u} \in V \ \wedge \ (\langle \boldsymbol{u}, \boldsymbol{v} \rangle = 0) \ \wedge \ \forall \boldsymbol{v} : (\boldsymbol{v} \in S)\}$$

$S^{\perp}$ is easily seen to be a subspace.

If $\dim(v) < \infty$, then $V = S \oplus S^{\perp}$.

Note: If $\dim(v) = \infty$, then in order to have $V = S \oplus S^{\perp}$ we require V to be a **Hilbert Space**.

Linear Transformations

Loosely speaking, a linear transformation is a mapping from one vector space to another that **preserves** vector space operations.

More precisely, let V, W be vector spaces over the same field F. A **linear transformation** is a mapping $T : V \to W$ such that

$$T(a\boldsymbol{u} + b\boldsymbol{v}) = aT(\boldsymbol{u}) + bT(\boldsymbol{v})$$

for all $a \in F$, $b \in F$ and $\boldsymbol{u} \in V$, $\boldsymbol{v} \in V$.

In this class we will be concerned with linear transformations between (real or complex) **Euclidean spaces**, or subspaces thereof.

Image

$$\text{image}\,(T) = \{\boldsymbol{w} | \, \boldsymbol{w} \in W \ \wedge \ T(\boldsymbol{v}) = \boldsymbol{w} \text{for some} \boldsymbol{v}\}$$

Nullspace

Also known as the kernel:

$$\ker(T) = \{\boldsymbol{v} | \, \boldsymbol{v} \in V \ \wedge \ (T(\boldsymbol{v}) = \boldsymbol{0})\}$$

Both the image and the nullspace are easily seen to be subspaces.

Rank

$$\operatorname{rank}(T) = \dim(\text{image } (T))$$

Nullity

$$\operatorname{null}(T) = \dim(\ker(T))$$

Rank plus nullity theorem

$$\operatorname{rank}(T) + \operatorname{null}(T) = \dim(V)$$

Matrices

Every linear transformation T has a **matrix representation**. If $T : \mathbb{E}^N \to \mathbb{E}^M$, $\mathbb{E} = \mathbb{R}$ or $\mathbb{C}$, then T is represented by an $M \times N$ matrix

$$A = \begin{pmatrix} a_{11} & \dots & a_{1N} \\ \vdots & \ddots & \vdots \\ a_{M1} & \dots & a_{MN} \end{pmatrix}$$

where $(a_{1i} \dots a_{Mi})^T = T(e_i)$ and $e_i = (0 \dots 1 \dots 0)^T$ is the i^{th} **standard basis** vector.

Note:A linear transformation can be represented with respect to any bases of $\mathbb{E}^N$ and $\mathbb{E}^M$, leading to a different A. We will always represent a linear transformation using the standard bases.

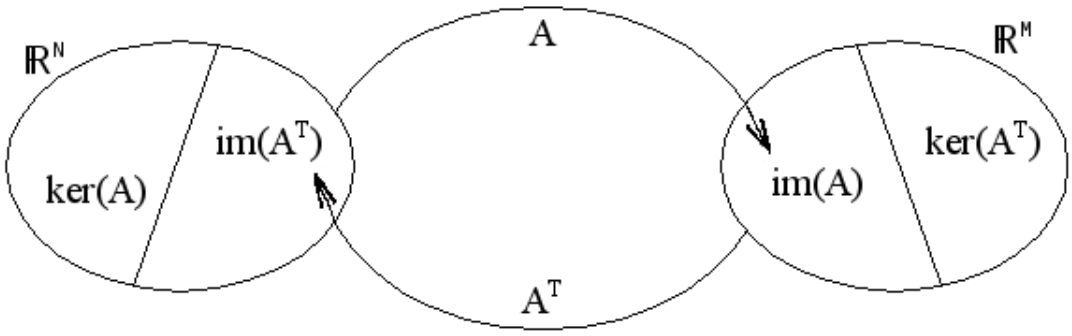
Column span

$$\text{colspan}\,(A) =< A >=\text{image}\,(A)$$

Duality

If $A : \mathbb{R}^N \to \mathbb{R}^M$, then

$$\ker^{\perp}(A) =\text{image}\,\left(A^T\right)$$



If $A : \mathbb{C}^N \to \mathbb{C}^M$, then

$$\ker^{\perp}(A) =\text{image}\,\left(A^{\text{H}}\right)$$

Inverses

The linear transformation/matrix A is **invertible** if and only if there exists a matrix B such that $AB = BA = I$ (identity).

Only **square** matrices can be invertible.

Let $A : \mathbb{F}^N \to \mathbb{F}^N$ be linear, $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$. The following are equivalent:

1. A is invertible (nonsingular)
2. $\text{rank}\,(A) = N$
3. $\text{null}\,(A) = 0$

4. $\det A \neq 0$
5. The columns of A form a basis.

If $A^{-1} = A^T$ (or A^{H} in the complex case), we say A is **orthogonal** (or **unitary**).

Signal Power

An interesting question you could ask about a signal is its average **power**. A signal's instantaneous power is defined to be its square, as if it were a voltage or current passing through a 1 Ω resistor. The average power is the average of the instantaneous power over some time interval. For a periodic signal, the natural time interval is clearly its period; for nonperiodic signals, a better choice would be entire time or time from onset. For a periodic signal, the average power is the square of the root-mean-squared (rms) value. We define the rms value of a periodic signal to be

Equation:

$$\mathrm{rms}\,(s) = \sqrt{2\frac{1}{T}\int_0^T s^2(t)\,\mathrm{d}\,t}$$

and thus its average power is $\mathrm{rms}^2(s)$.

Equation:

$$\begin{aligned}\mathrm{power}\,(s) &= \mathrm{rms}^2(s)\\ &= \frac{1}{T}\int_0^T s^2(t)\,\mathrm{d}\,t\end{aligned}$$

Exercise:

Problem: What is the rms value of the half-wave rectified sinusoid?

Solution:

A half-wave rectified sinusoid has half the average power of the original sine wave since it is zero half the time. A sine wave's average power equals $\frac{A^2}{2}$, making the rms value of the half-wave rectified signal $\frac{A}{2}$.

To find the average power in the frequency domain, we need to substitute the spectral representation of the signal into this expression.

Equation:

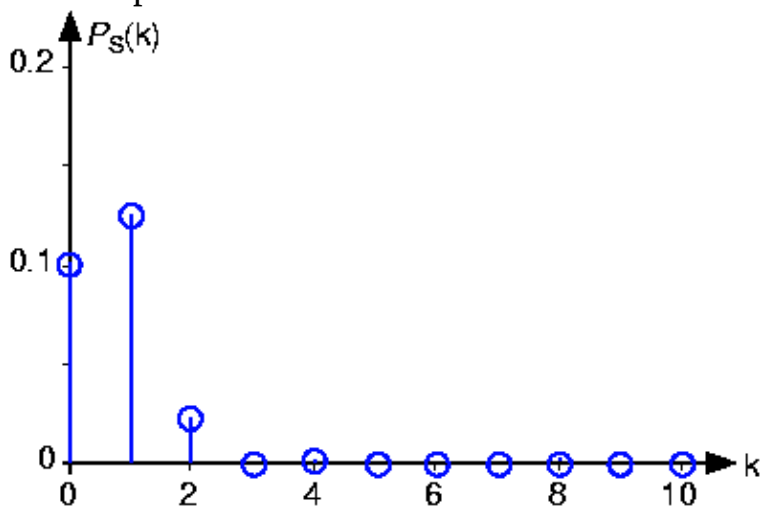
$$\text{power}\,(s) = \frac{1}{T}\int_0^T \left(a_0 + \sum_{k\ 1}^{\infty} a_k \cos\ \frac{2\pi kt}{T}\ + \sum_{k\ 1}^{\infty} b_k \sin\ \frac{2\pi kt}{T}\ \right)^2 \mathrm{d}\ t$$

The square inside the integral will contain all possible pairwise products. However, the orthogonality properties say that most of these crossterms integrate to zero. The survivors leave a rather simple expression for the power we seek.

Equation:

$$\text{power}\,(s) = {a_0}^2 + \frac{1}{2}\sum_{k\ 1}^{\infty} {a_k}^2 + {b_k}^2$$

Power Spectrum of a Half-Wave Rectified Sinusoid



Power spectrum of a half-wave rectified sinusoid.

It could well be that computing this sum is easier than integrating the signal's square. Furthermore, the contribution of each term in the Fourier series toward representing the signal can be measured by its contribution to the signal's average power. Thus, the power contained in a signal at its k th harmonic is $\frac{{a_k}^2+{b_k}^2}{2}$. The **power spectrum** $P_s(k)$, such as shown in [link], plots each harmonic's contribution to the total power.

Exercise:

Problem:

In stereophonic systems, deviation of a sine wave from the ideal is measured by the total harmonic distortion, which equals the total power in the harmonics higher than the first compared to power in the fundamental. Find an expression for the total harmonic distortion for any periodic signal. Is this calculation most easily performed in the time or frequency domain?

Solution:

Total harmonic distortion equals $\frac{\sum_{k=2}^{\infty} {a_k}^2+{b_k}^2}{{a_1}^2+{b_1}^2}$. Clearly, this quantity is most easily computed in the frequency domain. However, the numerator equals the the square of the signal's rms value minus the power in the average and the power in the first harmonic.

Signal Energy vs. Signal Power

The idea of the "size" of a signal is crucial to many applications. It is nice to know how much electricity can be used in a defibrillator without ill effects, for instance. It is also nice to know if the signal driving a set of headpones is enough to create a sound. While both of these examples deal with electric signals, they are clearly very different signals with very different tolerances. For this reason, it is convenient to quantify this idea of "size". This leads to the ideas of signal energy and signal power.

Signal Energy

Since we often think of signal as a function of varying amplitude through time, it seems to reason that a good measurement of the strength of a signal would be the area under the curve. However, this area may have a negative part. This negative part does not have less strength than a positive signal of the same size (reversing your grip on the paper clip in the socket is not going to make you any more lively). This suggests either squaring the signal or taking its absolute value, then finding the area under that curve. It turns out that what we call the **energy** of a signal is the area under the squared signal.

The energy of this signal is the shaded region.

Equation:

$$E_f = \int_{-\infty}^{\infty} (|f(t)|)^2 \, \mathrm{d}\, t$$

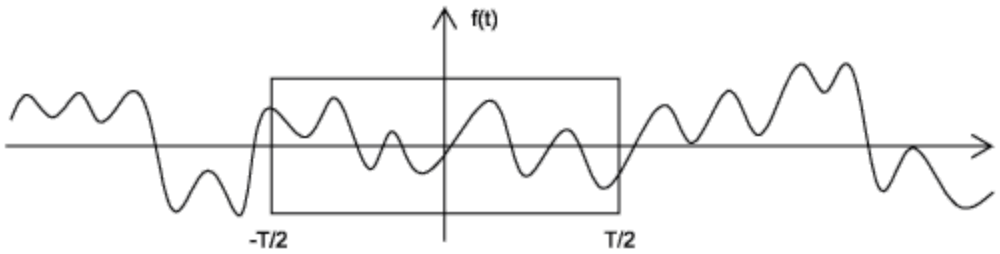
Signal Power

Our definition of energy seems reasonable, and it is. However, what if the signal does not decay? In this case we have infinite energy for any such signal. Does this mean that a sixty hertz sine wave feeding into your headphones is as strong as the sixty hertz sine wave coming out of your outlet? Obviously not. This is what leads us to the idea of **signal power**.

A simple, common signal with infinite energy.

Power is a time average of energy (energy per unit time). This is useful when the energy of the signal goes to infinity.
Equation:

$$P_f = \lim_{T \to \infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} \left(|f(t)| \right)^2 \, \mathrm{d}\, t$$



We compute the energy per a specific unit of time, then allow that time to go to infinity.

1. Compute $\frac{\text{Energy}}{T}$
2. Then look at $\lim_{T \to \infty} \frac{\text{Energy}}{T}$

P_f is often called the mean-square value of f. $\sqrt{P_f}$ is then called the **root mean squared** (**RMS**) value of f.

Energy vs. Power

- "Energy signals" have finite energy.
- "Power signals" have finite and non-zero power.

Exercise:

Problem: Are all energy signals also power signals?

Solution:

No. In fact, any signal with finite energy will have zero power.

Exercise:

Problem: Are all power signals also energy signals?

Solution:

No, any signal with non-zero power will have infinite energy.

Exercise:

Problem: Are **all** signals either energy or power signals?

Solution:

No. Any infinite-duration, increasing-magnitude function will not be either. (eg $f(t) = t$ is neither)

Introduction to Statistical Signal Processing

Digital Signal Processing

- Digital ≡ sampled, discrete-time, quantized
- Signal ≡ waveform, sequnce of measurements or observations
- Processing ≡ analyze, modify, filter, synthesize

Examples of Digital Signals

- sampled speech waveform
- "pixelized" image
- Dow-Jones Index

DSP Applications

- Filtering (noise reduction)
- Pattern recognition (speech, faces, fingerprints)
- Compression

A Major Difficulty

In many (perhaps most) DSP applications we don't have complete or perfect knowledge of the signals we wish to process. We are faced with many **unknowns** and **uncertainties**.

Examples

- noisy measurements
- unknown signal parameters
- noisy system or environmental conditions
- natural variability in the signals encountered

Functional Magnetic Resonance Imaging
[missing_resource: FMRI.png]

Challenges are measurement noise and intrinsic uncertainties in signal behavior.

How can we design signal processing algorithms in the face of such uncertainty?

Can we model the uncertainty and incorporate this model into the design process?

Statistical signal processing is the study of these questions.

Modeling Uncertainty

The most widely accepted and commonly used approach to modeling uncertainty is **Probability Theory** (although other alternatives exist such as Fuzzy Logic).

Probability Theory models uncertainty by specifying the **chance** of observing certain signals.

Alternatively, one can view probability as specifying the degree to which we **believe** a signal reflects the true **state of nature**.

Examples of Probabilistic Models

- errors in a measurement (due to an imprecise measuring device) modeled as realizations of a Gaussian random variable.
- uncertainty in the phase of a sinusoidal signal modeled as a uniform random variable on $[0, 2\pi)$.
- uncertainty in the number of photons stiking a CCD per unit time modeled as a Poisson random variable.

Statistical Inference

A **statistic** is a function of observed data.

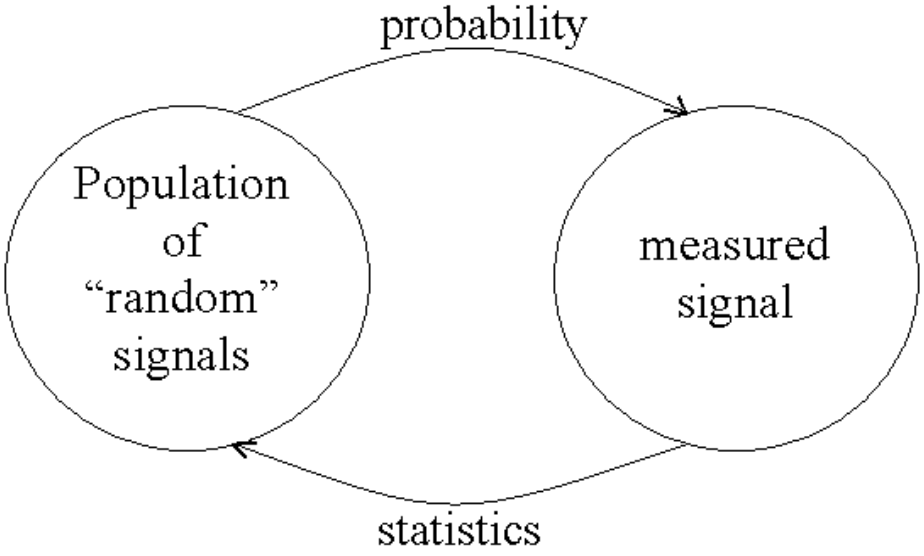
Example:
Suppose we observe N scalar values $x_1, \ldots, x_N$. The following are statistics:

- $\bar{x} = \frac{1}{N} \sum_{n=1}^{N} x_n$ (sample mean)
- $x_1, \ldots, x_N$ (the data itself)
- $\min \{x_1, \ldots, x_N\}$ (order statistic)
- $({x_1}^2 - x_2 \sin(x_3), e^{-(x_1 x_3)})$

A statistic **cannot** depend on **unknown parameters**.

Probability is used to model uncertainty.

Statistics are used to draw conclusions about probability models.



Probability models our uncertainty about signals we **may** observe.

Statistics reasons from the measured signal to the population of possible signals.

Statistical Signal Processing

- **Step 1** Postulate a probability model (or models) that reasonably capture the uncertainties at hand
- **Step 2** Collect data
- **Step 3** Formulate statistics that allow us to interpret or understand our probability model(s)

In this class

The two major kinds of problems that we will study are **detection** and **estimation**. Most SSP problems fall under one of these two headings.

Detection Theory

Given two (or more) probability models, which on best explains the signal?

Examples

1. Decode wireless comm signal into string of 0's and 1's
2. Pattern recognition
    - voice recognition
    - face recognition
    - handwritten character recognition
3. Anomaly detection
    - radar, sonar
    - irregular, heartbeat
    - gamma-ray burst in deep space

Estimation Theory

If our probability model has free parameters, what are the best parameter settings to describe the signal we've observed?

Examples

1. Noise reduction
2. Determine parameters of a sinusoid (phase, amplitude, frequency)
3. Adaptive filtering
    - track trajectories of space-craft
    - automatic control systems
    - channel equalization
4. Determine location of a submarine (sonar)
5. Seismology: estimate depth below ground of an oil deposit

Example:

Detection Example
Suppose we observe N tosses of an unfair coin. We would like to decide which side the coin favors, heads or tails.

- **Step 1** Assume each toss is a realization of a Bernoulli random variable.

$$\Pr[\text{Heads}] = p = 1 - \Pr[\text{Tails}]$$

  Must decide $p = \frac{1}{4}$ vs. $p = \frac{3}{4}$.
- **Step 2** Collect data $x_1, \ldots, x_N$

$$x_i = 1 \equiv \text{Heads}$$

$$x_i = 0 \equiv \text{Tails}$$

- **Step 3** Formulate a useful statistic

$$k = \sum_{n=1}^{N} x_n$$

  If $k < \frac{N}{2}$, guess $p = \frac{1}{4}$. If $k > \frac{N}{2}$, guess $p = \frac{3}{4}$.

Example:
Estimation Example
Suppose we take N measurements of a DC voltage A with a noisy voltmeter. We would like to estimate A.

- **Step 1** Assume a Gaussian noise model

$$x_n = A + w_n$$

  where $w_n \sim \mathcal{N}(0, 1)$.
- **Step 2** Gather data $x_1, \ldots, x_N$
- **Step 3** Compute the sample mean,

$$\widehat{A} = \frac{1}{N} \sum_{n=1}^{N} x_n$$

and use this as an estimate.

In these examples ([link] and [link]), we solved detection and estimation problems using intuition and heuristics (in Step 3).

This course will focus on developing principled and mathematically rigorous approaches to detection and estimation, using the theoretical framework of probability and statistics.

Summary

- DSP ≡ processing signals with computer algorithms.
- SSP ≡ statistical DSP ≡ processing in the presence of uncertainties and unknowns.

Introduction to Random Signals and Processes

Before now, you have probably dealt strictly with the theory behind signals and systems, as well as look at some the basic characteristics of signals and systems. In doing so you have developed an important foundation; however, most electrical engineers do not get to work in this type of fantasy world. In many cases the signals of interest are very complex due to the randomness of the world around them, which leaves them noisy and often corrupted. This often causes the information contained in the signal to be hidden and distorted. For this reason, it is important to understand these random signals and how to recover the necessary information.

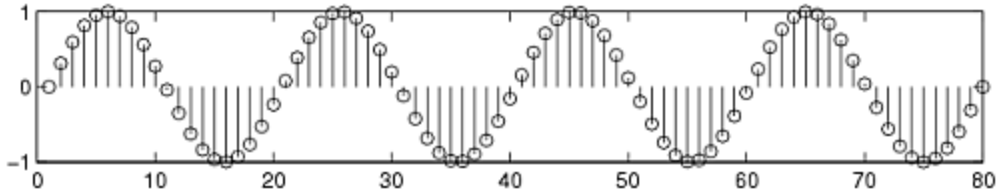
Signals: Deterministic vs. Stochastic

For this study of signals and systems, we will divide signals into two groups: those that have a fixed behavior and those that change randomly. As most of you have probably already dealt with the first type, we will focus on introducing you to random signals. Also, note that we will be dealing strictly with discrete-time signals since they are the signals we deal with in DSP and most real-world computations, but these same ideas apply to continuous-time signals.

Deterministic Signals

Most introductions to signals and systems deal strictly with **deterministic signals**. Each value of these signals are fixed and can be determined by a mathematical expression, rule, or table. Because of this, future values of any deterministic signal can be calculated from past values. For this reason, these signals are relatively easy to analyze as they do not change, and we can make accurate assumptions about their past and future behavior.

Deterministic Signal

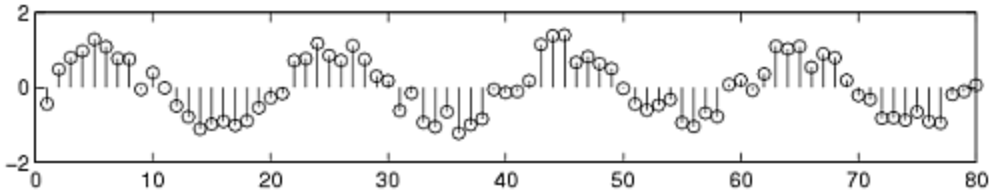


An example of a deterministic signal, the sine wave.

Stochastic Signals

Unlike deterministic signals, **stochastic signals**, or **random signals**, are not so nice. Random signals cannot be characterized by a simple, well-defined mathematical equation and their future values cannot be predicted. Rather, we must use probability and statistics to analyze their behavior. Also, because of their randomness, average values from a collection of signals are usually studied rather than analyzing one individual signal.

Random Signal



We have taken the above sine wave and added random noise to it to come up with a noisy, or random, signal. These are the types of signals that we wish to learn how to deal with so that we can recover the original sine wave.

Random Process

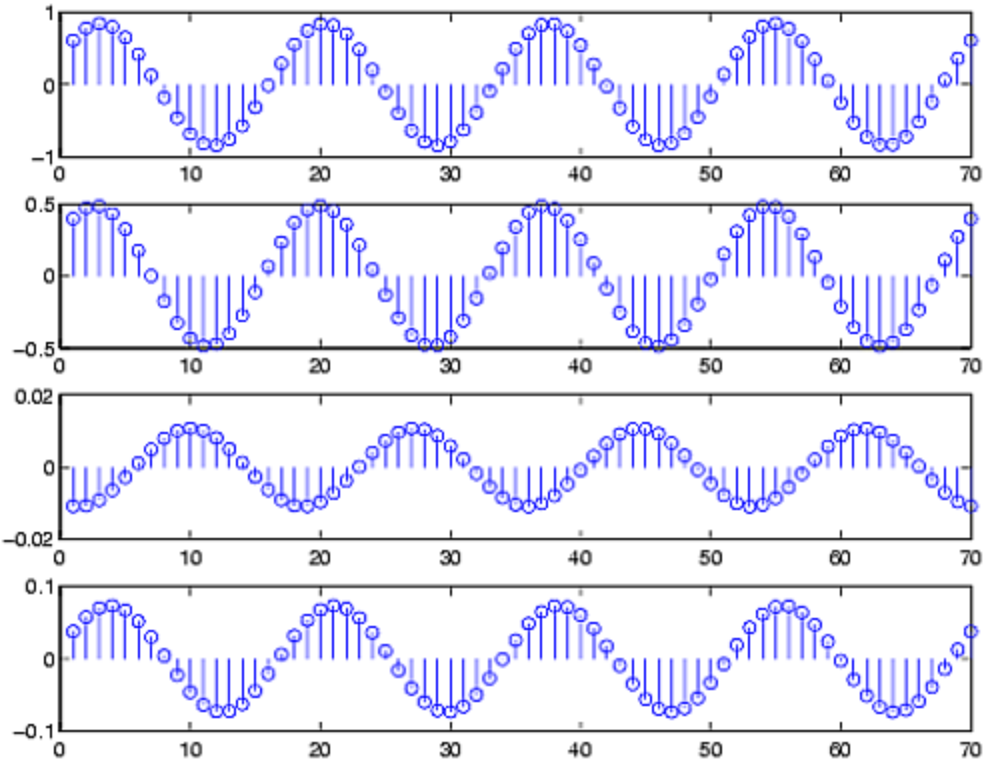
As mentioned above, in order to study random signals, we want to look at a collection of these signals rather than just one instance of that signal. This collection of signals is called a **random process**.

random process

A family or ensemble of signals that correspond to every possible outcome of a certain signal measurement. Each signal in this collection is referred to as a **realization** or **sample function** of the process.

Example:
As an example of a random process, let us look at the Random Sinusoidal Process below. We use $f[n] = A\sin(\omega n + \varphi)$ to represent the sinusoid with a given amplitude and phase. Note that the phase and amplitude of each sinusoid is based on a random number, thus making this a random process.

Random Sinusoidal Process



A random sinusoidal process, with the amplitude and phase being random numbers.

A random process is usually denoted by $X(t)$ or $X[n]$, with $x(t)$ or $x[n]$ used to represent an individual signal or waveform from this process.

In many notes and books, you might see the following notation and terms used to describe different types of random processes. For a **discrete random process**, sometimes just called a **random sequence**, t represents time that has a finite number of values. If t can take on any value of time, we have a **continuous random process**. Often times discrete and continuous refer to the amplitude of the process, and process or sequence refer to the nature of the time variable. For this study, we often just use **random process** to refer to a general collection of discrete-time signals, as seen above in [link].

Introduction to Stochastic Processes

Definitions, distributions, and stationarity

Stochastic Process

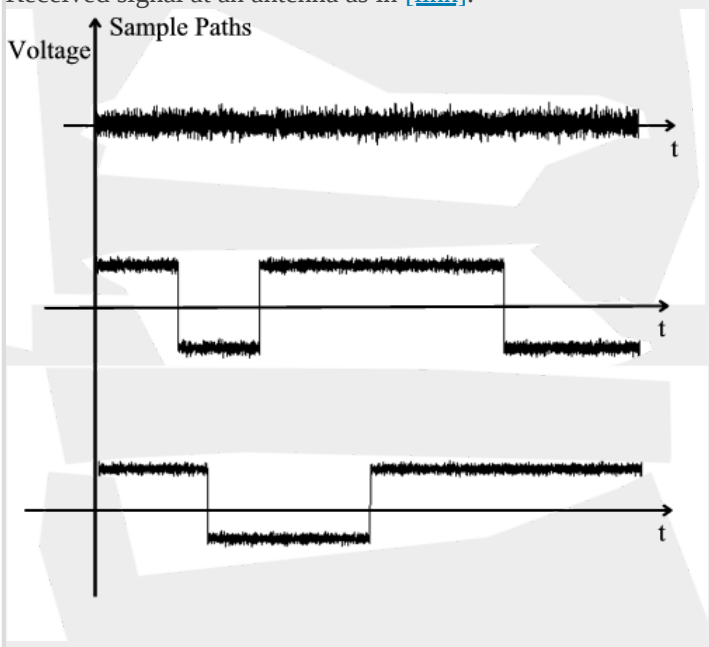
Given a sample space, a stochastic process is an indexed collection of random variables defined for each $\omega \in \Omega$.

Equation:

$$\forall t, t \in \mathbb{R} : (X_t(\omega))$$

Example:

Received signal at an antenna as in [link].



For a given t, $X_t(\omega)$ is a random variable with a distribution

Equation:

First-order distribution

$$\begin{aligned} F_{X_t}(b) &= \Pr[X_t \leq b] \\ &= \Pr[\{\omega \in \Omega | \, X_t(\omega) \leq b\}] \end{aligned}$$

First-order stationary process

If $F_{X_t}(b)$ is not a function of time then X_t is called a first-order stationary process.

Equation:

Second-order distribution

$$F_{X_{t_1},X_{t_2}}(b_1, b_2) = \Pr[X_{t_1} \leq b_1, X_{t_2} \leq b_2]$$

for all $t_1 \in \mathbb{R}$, $t_2 \in \mathbb{R}$, $b_1 \in \mathbb{R}$, $b_2 \in \mathbb{R}$
Equation:

Nth-order distribution

$$F_{X_{t_1},X_{t_2},\ldots,X_{t_N}}(b_1,b_2,\ldots,b_N) = \Pr[X_{t_1} \leq b_1,\ldots,X_{t_N} \leq b_N]$$

Nth-order stationary : A random process is stationary of order N if
Equation:

$$F_{X_{t_1},X_{t_2},\ldots,X_{t_N}}(b_1,b_2,\ldots,b_N) = F_{X_{t_1+T},X_{t_2+T},\ldots,X_{t_N+T}}(b_1,b_2,\ldots,b_N)$$

Strictly stationary : A process is strictly stationary if it is Nth order stationary for all N.

Example:
$X_t = \cos(2\pi f_0 t + \Theta(\omega))$ where f_0 is the deterministic carrier frequency and $\Theta(\omega) : \Omega \to \mathbb{R}$ is a random variable defined over $[-\pi, \pi]$ and is assumed to be a uniform random variable; i.e.,
$f_\Theta(\theta) = \begin{cases} \frac{1}{2\pi} & \text{if } \theta \in [-\pi,\pi] \\ 0 & \text{otherwise} \end{cases}$
Equation:

$$\begin{aligned} F_{X_t}(b) &= \Pr[X_t \leq b] \\ &= \Pr[\cos(2\pi f_0 t + \Theta) \leq b] \end{aligned}$$

Equation:

$$F_{X_t}(b) = \Pr[-\pi \leq 2\pi f_0 t + \Theta \leq -\arccos(b)] + \Pr[\arccos(b) \leq 2\pi f_0 t + \Theta \leq \pi]$$

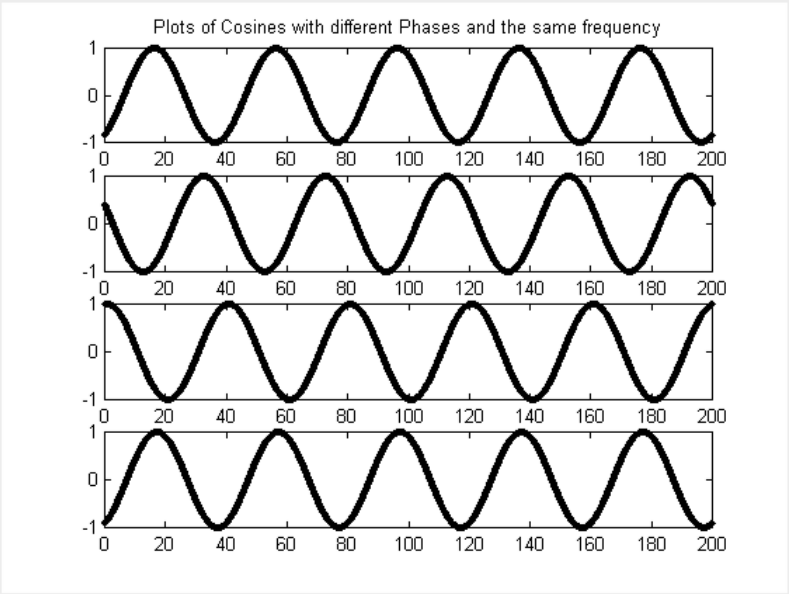
Equation:

$$\begin{aligned} F_{X_t}(b) &= \int_{(-\pi)-2\pi f_0 t}^{(-\arccos(b))-2\pi f_0 t} \tfrac{1}{2\pi}\, \mathrm{d}\,\theta + \int_{\arccos(b)-2\pi f_0 t}^{\pi-2\pi f_0 t} \tfrac{1}{2\pi}\, \mathrm{d}\,\theta \\ &= (2\pi - 2\arccos(b))\tfrac{1}{2\pi} \end{aligned}$$

Equation:

$$\begin{aligned} f_{X_t}(x) &= \tfrac{\mathrm{d}}{\mathrm{d}x}\left(1 - \tfrac{1}{\pi}\arccos(x)\right) \\ &= \begin{cases} \frac{1}{\pi\sqrt{1-x^2}} & \text{if } |x| \leq 1 \\ 0 & \text{otherwise} \end{cases} \end{aligned}$$

This process is stationary of order 1.



The second order stationarity can be determined by first considering conditional densities and the joint density. Recall that

Equation:

$$X_t = \cos(2\pi f_0 t + \Theta)$$

Then the relevant step is to find

Equation:

$$\Pr[X_{t_2} \leq b_2 \mid X_{t_1} = x_1]$$

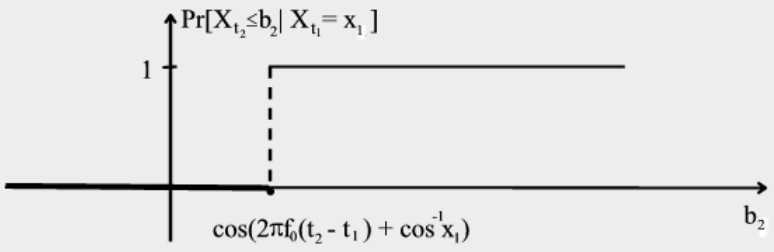
Note that

Equation:

$$(X_{t_1} = x_1 = \cos(2\pi f_0 t + \Theta)) \Rightarrow (\Theta = \arccos(x_1) - 2\pi f_0 t)$$

Equation:

$$\begin{aligned} X_{t_2} &= \cos(2\pi f_0 t_2 + \arccos(x_1) - 2\pi f_0 t_1) \\ &= \cos(2\pi f_0 \left(t_2 - t_1\right) + \arccos(x_1)) \end{aligned}$$

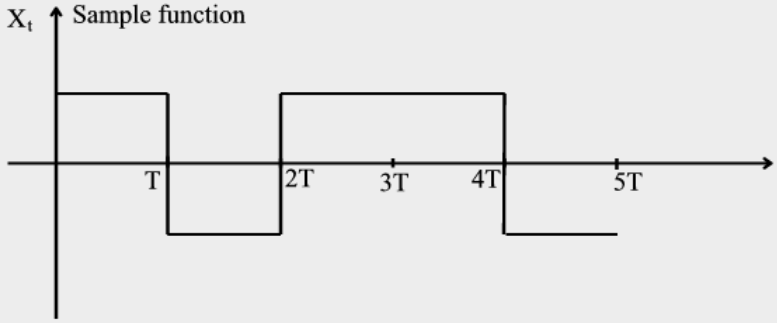


Equation:

$$F_{X_{t_2},X_{t_1}}(b_2,b_1) = \int_{-\infty}^{b_1} f_{X_{t_1}}(x_1)\Pr[X_{t_2} \leq b_2 \mid X_{t_1} = x_1]\ \mathrm{d}\ x_1$$

Note that this is only a function of $t_2 - t_1$.

Example:
Every T seconds, a fair coin is tossed. If heads, then $X_t = 1$ for $nT \leq t < (n+1)T$. If tails, then $X_t = -1$ for $nT \leq t < (n+1)T$.



Equation:

$$p_{X_t}(x) = \begin{cases} \frac{1}{2} & \text{if } x = 1 \\ \frac{1}{2} & \text{if } x = -1 \end{cases}$$

for all $t \in \mathbb{R}$. X_t is stationary of order 1.
Second order probability mass function
Equation:

$$p_{X_{t_1}X_{t_2}}(x_1,x_2) = p_{X_{t_2}|X_{t_1}}(x_2|x_1)p_{X_{t_1}}(x_1)$$

The conditional pmf
Equation:

$$p_{X_{t_2}|X_{t_1}}(x_2|x_1) = \begin{cases} 0 & \text{if } x_2 \neq x_1 \\ 1 & \text{if } x_2 = x_1 \end{cases}$$

when $nT \leq t_1 < (n+1)T$ and $nT \leq t_2 < (n+1)T$ for some n.
Equation:

$$p_{X_{t_2}|X_{t_1}}(x_2|x_1) = p_{X_{t_2}}(x_2)$$

for all x_1 and for all x_2 when $nT \leq t_1 < (n+1)T$ and $mT \leq t_2 < (m+1)T$ with $n \neq m$
Equation:

$$p_{X_{t_2}X_{t_1}}(x_2,x_1) = \begin{cases} 0 & \text{if } x_2 \neq x_1 \text{for } nT \leq t_1, t_2 < (n+1)T \\ p_{X_{t_1}}(x_1) & \text{if } x_2 = x_1 \text{for } nT \leq t_1, t_2 < (n+1)T \\ p_{X_{t_1}}(x_1)p_{X_{t_2}}(x_2) & \text{if } n \neq m \text{for } (nT \leq t_1 < (n+1)T) \ \wedge \ (mT \leq t_2 < (m+1)T) \end{cases}$$

The Gaussian Process

A random process $X(t)$ is Gaussian if the joint density of the N amplitudes $\{X(t_1), \ldots,, X(t_N)\}$ comprise a Gaussian random vector. The elements of the required covariance matrix equal the covariance between the appropriate amplitudes: $K_{i,j} = K_X(t_i, t_j)$. Assuming the mean is known, the entire structure of the Gaussian random process is specified once the correlation function or, equivalently, the power spectrum is known. As linear transformations of Gaussian random processes yield another Gaussian process, linear operations such as differentiation, integration, linear filtering, sampling, and summation with other Gaussian processes result in a Gaussian process.

Sampling and Random Sequences

The usual Sampling Theorem applies to random processes, with the spectrum of interest beign the power spectrum. If stationary process $X(t)$ is bandlimited - $\mathscr{S}_X(\omega) = 0$, $|\omega| > W$, as long as the sampling interval T satisfies the classic constraint $T < \frac{\pi}{W}$ the sequence $X(lT)$ represents the original process. A sampled process is itself a random process defined over discrete time. Hence, all of the random process notions introduced in the previous section apply to the random sequence $\widetilde{X}(l) \equiv X(lT)$. The correlation functions of these two processes are related as

$$R_{\widetilde{X}}(k) = E\left[\widetilde{X}(l)\widetilde{X}(l+k)\right] = R_X(kT)$$

We note especially that for distinct samples of a random process to be uncorrelated, the correlation function $R_X(kT)$ must equal zero for all non-zero k. This requirement places severe restrictions on the correlation function (hence the power spectrum) of the original process. One correlation function satisfying this property is derived from the random process which has a bandlimited, constant-valued power spectrum over precisely the frequency region needed to satisfy the sampling criterion. **No other power spectrum satisfying the sampling criterion has this property**. Hence, sampling does not normally yield uncorrelated amplitudes, meaning that **discrete-time white noise** is a rarity. White noise has a correlation function given by $R_{\widetilde{X}}(k) = \sigma^2\delta(k)$, where $\delta(\cdot)$ is the unit sample. The power spectrum of white noise is a constant: $\mathscr{S}_{\widetilde{X}}(\omega) = \sigma^2$.

Stationary and Nonstationary Random Processes

Introduction

From the definition of a random process, we know that all random processes are composed of random variables, each at its own unique point in time. Because of this, random processes have all the properties of random variables, such as mean, correlation, variances, etc.. When dealing with groups of signals or sequences it will be important for us to be able to show whether of not these statistical properties hold true for the entire random process. To do this, the concept of **stationary processes** has been developed. The general definition of a stationary process is:

stationary process
: a random process where all of its statistical properties do not vary with time

Processes whose statistical properties do change are referred to as **nonstationary**.

Understanding the basic idea of stationarity will help you to be able to follow the more concrete and mathematical definition to follow. Also, we will look at various levels of stationarity used to describe the various types of stationarity characteristics a random process can have.

Distribution and Density Functions

In order to properly define what it means to be stationary from a mathematical standpoint, one needs to be somewhat familiar with the concepts of distribution and density functions. If you can remember your statistics then feel free to skip this section!

Recall that when dealing with a single random variable, the **probability distribution function** is a simply tool used to identify the probability that our observed random variable will be less than or equal to a given number. More precisely, let X be our random variable, and let x be our given value; from this we can define the distribution function as

Equation:

$$F_x(x) = \Pr[X \leq x]$$

This same idea can be applied to instances where we have multiple random variables as well. There may be situations where we want to look at the probability of event X **and** Y both occurring. For example, below is an example of a second-order **joint distribution function**.
Equation:

$$F_x(x, y) = \Pr[X \leq x, Y \leq y]$$

While the distribution function provides us with a full view of our variable or processes probability, it is not always the most useful for calculations. Often times we will want to look at its derivative, the **probability density function (pdf)**. We define the the pdf as
Equation:

$$f_x(x) = \frac{\mathrm{d}}{\mathrm{d}\, x} F_x(x)$$

Equation:

$$f_x(x)\,\mathrm{dx} = \Pr[x < X \leq x + \mathrm{dx}]$$

[link] reveals some of the physical significance of the density function. This equations tells us the probability that our random variable falls within a given interval can be approximated by $f_x(x)$ dx. From the pdf, we can now use our knowledge of integrals to evaluate probabilities from the above approximation. Again we can also define a **joint density function** which will include multiple random variables just as was done for the distribution function. The density function is used for a variety of calculations, such as finding the expected value or proving a random variable is stationary, to name a few.

Note: The above examples explain the distribution and density functions in terms of a single random variable, X. When we are dealing with signals and random processes, remember that we will have a set of random variables where a different random variable will occur at each time instance of the random process, $X(t_k)$. In other words, the distribution and density function will also need to take into account the choice of time.

Stationarity

Below we will now look at a more in depth and mathematical definition of a stationary process. As was mentioned previously, various levels of stationarity exist and we will look at the most common types.

First-Order Stationary Process

A random process is classified as **first-order stationary** if its first-order probability density function remains equal regardless of any shift in time to its time origin. If we let x_{t_1} represent a given value at time t_1, then we define a first-order stationary as one that satisfies the following equation:
Equation:

$$f_x(x_{t_1}) = f_x(x_{t_1+\tau})$$

The physical significance of this equation is that our density function, $f_x(x_{t_1})$, is completely independent of t_1 and thus any time shift, τ.

The most important result of this statement, and the identifying characteristic of any first-order stationary process, is the fact that the mean is a constant, independent of any time shift. Below we show the results for a random process, X, that is a discrete-time signal, $x[n]$.
Equation:

$$\begin{aligned} X &= m_x[n] \\ &= E[x[n]] \\ &= \text{constant (independent of n)} \end{aligned}$$

Second-Order and Strict-Sense Stationary Process

A random process is classified as **second-order stationary** if its second-order probability density function does not vary over any time shift applied to both values. In other words, for values x_{t_1} and x_{t_2} then we will have the following be equal for an arbitrary time shift τ.
Equation:

$$f_x(x_{t_1}, x_{t_2}) = f_x(x_{t_1+\tau}, x_{t_2+\tau})$$

From this equation we see that the absolute time does not affect our functions, rather it only really depends on the time difference between the two variables. Looked at another way, this equation can be described as
Equation:

$$\Pr[X(t_1) \leq x_1, X(t_2) \leq x_2] = \Pr[X(t_1 + \tau) \leq x_1, X(t_2 + \tau) \leq x_2]$$

These random processes are often referred to as **strict sense stationary (SSS)** when **all** of the distribution functions of the process are unchanged regardless of the time shift applied to them.

For a second-order stationary process, we need to look at the autocorrelation function to see its most important property. Since we have already stated that a second-order stationary process depends only on the time difference, then all of these types of processes have the following property:
Equation:

$$\begin{aligned} R_{xx}(t, t+\tau) &= E[X(t+\tau)] \\ &= R_{xx}(\tau) \end{aligned}$$

Wide-Sense Stationary Process

As you begin to work with random processes, it will become evident that the strict requirements of a SSS process is more than is often necessary in order to adequately approximate our calculations on random processes. We define a final type of stationarity, referred to as **wide-sense stationary (WSS)**, to have slightly more relaxed requirements but ones that are still enough to provide us with adequate results. In order to be WSS a random process only needs to meet the following two requirements.

1. $X = E[x[n]] = \text{constant}$
2. $E[X(t+\tau)] = R_{xx}(\tau)$

Note that a second-order (or SSS) stationary process will always be WSS; however, the reverse will not always hold true.

Correlation and Covariance of a Random Signal

When we take the expected value, or average, of a random process, we measure several important characteristics about how the process behaves in general. This proves to be a very important observation. However, suppose we have several random processes measuring different aspects of a system. The relationship between these different processes will also be an important observation. The covariance and correlation are two important tools in finding these relationships. Below we will go into more details as to what these words mean and how these tools are helpful. Note that much of the following discussions refer to just random variables, but keep in mind that these variables can represent random signals or random processes.

Covariance

To begin with, when dealing with more than one random process, it should be obvious that it would be nice to be able to have a number that could quickly give us an idea of how similar the processes are. To do this, we use the **covariance**, which is analogous to the variance of a single variable.

Covariance
: A measure of how much the deviations of two or more variables or processes match.

For two processes, X and Y, if they are **not** closely related then the covariance will be small, and if they are similar then the covariance will be large. Let us clarify this statement by describing what we mean by "related" and "similar." Two processes are "closely related" if their distribution spreads are almost equal and they are around the same, or a very slightly different, mean.

Mathematically, covariance is often written as σ_{xy} and is defined as
Equation:

$$\begin{aligned} \operatorname{cov}(X,Y) &= \sigma_{xy} \\ &= E\Big[\Big(X-X\Big)\Big(Y-Y\Big)\Big] \end{aligned}$$

This can also be reduced and rewritten in the following two forms:
Equation:

$$\sigma_{xy} = (xy) - \bar{x}\bar{y}$$

Equation:

$$\sigma_{xy} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left(X - X\right) \left(Y - Y\right) f(x, y) \, \mathrm{d}\, x \, \mathrm{d}\, y$$

Useful Properties

- If X and Y are independent and uncorrelated or one of them has zero mean value, then

$$\sigma_{xy} = 0$$

- If X and Y are orthogonal, then

$$\sigma_{xy} = -\left(E[X]E[Y]\right)$$

- The covariance is symmetric

$$\operatorname{cov}\left(X, Y\right) = \operatorname{cov}\left(Y, X\right)$$

- Basic covariance identity

$$\operatorname{cov}\left(X + Y, Z\right) = \operatorname{cov}\left(X, Z\right) + \operatorname{cov}\left(Y, Z\right)$$

- Covariance of equal variables

$$\operatorname{cov}\left(X, X\right) = \operatorname{Var}\left(X\right)$$

Correlation

For anyone who has any kind of statistical background, you should be able to see that the idea of dependence/independence among variables and signals plays an important role when dealing with random processes. Because of this, the **correlation** of two variables provides us with a measure of how the two variables affect one another.

Correlation

A measure of how much one random variable depends upon the other.

This measure of association between the variables will provide us with a clue as to how well the value of one variable can be predicted from the value of the other. The correlation is equal to the average of the product of two random variables and is defined as

Equation:

$$\begin{aligned} \operatorname{cor}(X,Y) &= E[XY] \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyf(x,y)\,\mathrm{d}\,x\,\mathrm{d}\,y \end{aligned}$$

Correlation Coefficient

It is often useful to express the correlation of random variables with a range of numbers, like a percentage. For a given set of variables, we use the **correlation coefficient** to give us the linear relationship between our variables. The correlation coefficient of two variables is defined in terms of their covariance and standard deviations, denoted by σ_x, as seen below

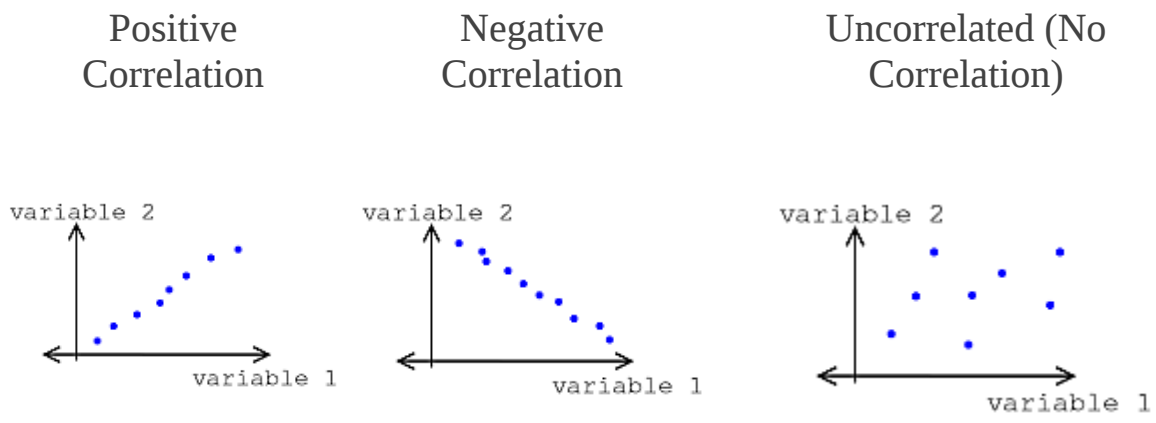
Equation:

$$\rho = \frac{\operatorname{cov}(X,Y)}{\sigma_x \sigma_y}$$

where we will always have

$$-1 \leq \rho \leq 1$$

This provides us with a quick and easy way to view the correlation between our variables. If there is no relationship between the variables then the correlation coefficient will be zero and if there is a perfect positive match it will be one. If there is a perfect inverse relationship, where one set of variables increases while the other decreases, then the correlation coefficient will be negative one. This type of correlation is often referred to more specifically as the **Pearson's Correlation Coefficient**,or Pearson's Product Moment Correlation.



Types of Correlation

Note: So far we have dealt with correlation simply as a number relating the relationship between any two variables. However, since our goal will be to relate random processes to each other, which deals with signals as a function of time, we will want to continue this study by looking at **correlation functions**.

Example

Now let us take just a second to look at a simple example that involves calculating the covariance and correlation of two sets of random numbers. We are given the following data sets:

$$x = \{3, 1, 6, 3, 4\}$$

$$y = \{1, 5, 3, 4, 3\}$$

To begin with, for the covariance we will need to find the expected value, or mean, of x and y.

$$\bar{x} = \frac{1}{5}\left(3 + 1 + 6 + 3 + 4\right) = 3.4$$

$$\bar{y} = \frac{1}{5}\left(1 + 5 + 3 + 4 + 3\right) = 3.2$$

$$xy = \frac{1}{5}\left(3 + 5 + 18 + 12 + 12\right) = 10$$

Next we will solve for the standard deviations of our two sets using the formula below (for a review click here).

$$\sigma = \sqrt{E\left[\left(X - E[X]\right)^2\right]}$$

$$\sigma_x = \sqrt{\frac{1}{5}\left(0.16 + 5.76 + 6.76 + 0.16 + 0.36\right)} = 1.625$$

$$\sigma_y = \sqrt{\frac{1}{6}\left(4.84 + 3.24 + 0.04 + 0.64 + 0.04\right)} = 1.327$$

Now we can finally calculate the covariance using one of the two formulas found above. Since we calculated the three means, we will use that formula since it will be much simpler.

$$\sigma_{xy} = 10 - 3.4 \times 3.2 = -0.88$$

And for our last calculation, we will solve for the correlation coefficient, ρ.

$$\rho = \frac{-0.88}{1.625 \times 1.327} = -0.408$$

Matlab Code for Example

The above example can be easily calculated using Matlab. Below I have included the code to find all of the values above.

```
x = [3 1 6 3 4];
y = [1 5 3 4 3];

mx = mean(x)
my = mean(y)
mxy = mean(x.*y)

% Standard Dev. from built-in Matlab
Functions
std(x,1)
std(y,1)

% Standard Dev. from Equation Above
(same result as std(?,1))
sqrt( 1/5 * sum((x-mx).^2))
sqrt( 1/5 * sum((y-my).^2))

cov(x,y,1)

corrcoef(x,y)
```

Autocorrelation of Random Processes

Before diving into a more complex statistical analysis of random signals and processes, let us quickly review the idea of correlation. Recall that the correlation of two signals or variables is the expected value of the product of those two variables. Since our focus will be to discover more about a random process, a collection of random signals, then imagine us dealing with two samples of a random process, where each sample is taken at a different point in time. Also recall that the key property of these random processes is that they are now functions of time; imagine them as a collection of signals. The expected value of the product of these two variables (or samples) will now depend on how quickly they change in regards to **time**. For example, if the two variables are taken from almost the same time period, then we should expect them to have a high correlation. We will now look at a correlation function that relates a pair of random variables from the same process to the time separations between them, where the argument to this correlation function will be the time difference. For the correlation of signals from two different random process, look at the crosscorrelation function.

Autocorrelation Function

The first of these correlation functions we will discuss is the **autocorrelation**, where each of the random variables we will deal with come from the same random process.

Autocorrelation
: the expected value of the product of a random variable or signal realization with a time-shifted version of itself

With a simple calculation and analysis of the autocorrelation function, we can discover a few important characteristics about our random process. These include:

1. How quickly our random signal or processes changes with respect to the time function

2. Whether our process has a periodic component and what the expected frequency might be

As was mentioned above, the autocorrelation function is simply the expected value of a product. Assume we have a pair of random variables from the same process, $X_1 = X(t_1)$ and $X_2 = X(t_2)$, then the autocorrelation is often written as
Equation:

$$\begin{aligned} R_{\text{xx}}(t_1, t_2) &= E[X_1 X_2] \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1 x_2 f(x_1, x_2) \, \mathrm{d}\, x_2 \, \mathrm{d}\, x_1 \end{aligned}$$

The above equation is valid for stationary and nonstationary random processes. For stationary processes, we can generalize this expression a little further. Given a wide-sense stationary processes, it can be proven that the expected values from our random process will be independent of the origin of our time function. Therefore, we can say that our autocorrelation function will depend on the time difference and not some absolute time. For this discussion, we will let $\tau = t_2 - t_1$, and thus we generalize our autocorrelation expression as
Equation:

$$\begin{aligned} R_{\text{xx}}(t, t+\tau) &= R_{\text{xx}}(\tau) \\ &= E[X(t) X(t+\tau)] \end{aligned}$$

for the continuous-time case. In most DSP course we will be more interested in dealing with real signal sequences, and thus we will want to look at the discrete-time case of the autocorrelation function. The formula below will prove to be more common and useful than [link]:
Equation:

$$R_{\text{xx}}[n, n+m] = \sum_{n=-\infty}^{\infty} x[n] x[n+m]$$

And again we can generalize the notation for our autocorrelation function as **Equation:**

$$\begin{aligned} R_{\text{xx}}[n, n+m] &= R_{\text{xx}}[m] \\ &= E[X[n]X[n+m]] \end{aligned}$$

Properties of Autocorrelation

Below we will look at several properties of the autocorrelation function that hold for **stationary** random processes.

- Autocorrelation is an even function for τ

$$R_{\text{xx}}(\tau) = R_{\text{xx}}(-\tau)$$

- The mean-square value can be found by evaluating the autocorrelation where $\tau = 0$, which gives us

$$R_{\text{xx}}(0) = X^2$$

- The autocorrelation function will have its largest value when $\tau = 0$. This value can appear again, for example in a periodic function at the values of the equivalent periodic points, but will never be exceeded.

$$R_{\text{xx}}(0) \geq |R_{\text{xx}}(\tau)|$$

- If we take the autocorrelation of a period function, then $R_{\text{xx}}(\tau)$ will also be periodic with the same frequency.

Estimating the Autocorrleation with Time-Averaging

Sometimes the whole random process is not available to us. In these cases, we would still like to be able to find out some of the characteristics of the

stationary random process, even if we just have part of one sample function. In order to do this we can **estimate** the autocorrelation from a given interval, 0 to T seconds, of the sample function.
Equation:

$$\check{R}_{\mathrm{xx}}(\tau) = \frac{1}{T-\tau} \int_0^{T-\tau} x(t)x(t+\tau)\ \mathrm{d}\,t$$

However, a lot of times we will not have sufficient information to build a complete continuous-time function of one of our random signals for the above analysis. If this is the case, we can treat the information we do know about the function as a discrete signal and use the discrete-time formula for estimating the autocorrelation.
Equation:

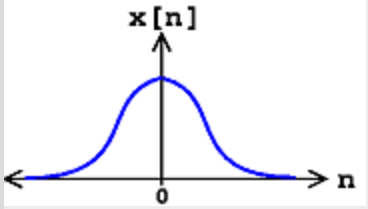
$$\check{R}_{\mathrm{xx}}[m] = \frac{1}{N-m} \sum_{n=0}^{N-m-1} x[n]x[n+m]$$

Examples

Below we will look at a variety of examples that use the autocorrelation function. We will begin with a simple example dealing with Gaussian White Noise (GWN) and a few basic statistical properties that will prove very useful in these and future calculations.

Example:
We will let $x[n]$ represent our GWN. For this problem, it is important to remember the following fact about the mean of a GWN function:

$$E[x[n]] = 0$$



Gaussian density function. By examination, can easily see that the above statement is true - the mean equals zero.

Along with being **zero-mean**, recall that GWN is always **independent**. With these two facts, we are now ready to do the short calculations required to find the autocorrelation.

$$R_{\text{xx}}[n, n+m] = E[x[n]x[n+m]]$$

Since the function, $x[n]$, is independent, then we can take the product of the individual expected values of both functions.

$$R_{\text{xx}}[n, n+m] = E[x[n]]E[x[n+m]]$$

Now, looking at the above equation we see that we can break it up further into two conditions: one when m and n are equal and one when they are not equal. When they are equal we can combine the expected values. We are left with the following piecewise function to solve:

$$R_{\text{xx}}[n, n+m] = \begin{array}{ll} E[x[n]]E[x[n+m]] & \text{if } \ m \neq 0 \\ E\left[x^2[n]\right] & \text{if } \ m = 0 \end{array}$$

We can now solve the two parts of the above equation. The first equation is easy to solve as we have already stated that the expected value of $x[n]$ will be zero. For the second part, you should recall from statistics that the

expected value of the square of a function is equal to the variance. Thus we get the following results for the autocorrelation:

$$R_{\text{xx}}[n, n+m] = \begin{matrix} 0 \text{ if } m \neq 0 \\ \sigma^2 \text{ if } m = 0 \end{matrix}$$

Or in a more concise way, we can represent the results as

$$R_{\text{xx}}[n, n+m] = \sigma^2 \delta[m]$$

Crosscorrelation of Random Processes

Before diving into a more complex statistical analysis of random signals and processes, let us quickly review the idea of correlation. Recall that the correlation of two signals or variables is the expected value of the product of those two variables. Since our main focus is to discover more about random processes, a collection of random signals, we will deal with two random processes in this discussion, where in this case we will deal with samples from two **different** random processes. We will analyze the expected value of the product of these two variables and how they correlate to one another, where the argument to this correlation function will be the time difference. For the correlation of signals from the same random process, look at the autocorrelation function.

Crosscorrelation Function

When dealing with multiple random processes, it is also important to be able to describe the relationship, if any, between the processes. For example, this may occur if more than one random signal is applied to a system. In order to do this, we use the **crosscorrelation function**, where the variables are instances from two different wide sense stationary random processes.

Crosscorrelation
: if two processes are wide sense stationary, the expected value of the product of a random variable from one random process with a time-shifted, random variable from a different random process

Looking at the generalized formula for the crosscorrelation, we will represent our two random processes by allowing $U = U(t)$ and $V = V(t - \tau)$. We will define the crosscorrelation function as
Equation:

$$\begin{aligned} R_{uv}(t, t-\tau) &= E[UV] \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} uvf(u,v) \, \mathrm{d}\, v \, \mathrm{d}\, u \end{aligned}$$

Just as the case with the autocorrelation function, if our input and output, denoted as $U(t)$ and $V(t)$, are at least jointly wide sense stationary, then the crosscorrelation does not depend on absolute time; it is just a function of the time difference. This means we can simplify our writing of the above function as

Equation:

$$R_{uv}(\tau) = E[UV]$$

or if we deal with two real signal sequences, $x[n]$ and $y[n]$, then we arrive at a more commonly seen formula for the discrete crosscorrelation function. See the formula below and notice the similarities between it and the convolution of two signals:

Equation:

$$\begin{aligned} R_{xy}(n, n-m) &= R_{xy}(m) \\ &= \sum_{n=-\infty}^{\infty} x[n]y[n-m] \end{aligned}$$

Properties of Crosscorrelation

Below we will look at several properties of the crosscorrelation function that hold for two **wide sense stationary (WSS)** random processes.

- Crosscorrelation is **not** an even function; however, it does have a unique symmetry property:
  Equation:

  $$R_{xy}(-\tau) = R_{yx}(\tau)$$

- The maximum value of the crosscorrelation is not always when the shift equals zero; however, we can prove the following property revealing to us what value the maximum cannot exceed.
  Equation:

$$|R_{xy}(\tau)| \leq \sqrt{R_{xx}(0)R_{yy}(0)}$$

- When two random processes are statistically independent then we have **Equation:**

$$R_{xy}(\tau) = R_{yx}(\tau)$$

Examples

Exercise:

Problem:

Let us begin by looking at a simple example showing the relationship between two sequences. Using [link], find the crosscorrelation of the sequences

$$x[n] = \{\ldots, 0, 0, 2, -3, 6, 1, 3, 0, 0, \ldots\}$$

$$y[n] = \{\ldots, 0, 0, 1, -2, 4, 1, -3, 0, 0, \ldots\}$$

for each of the following possible time shifts: $m = \{0, 3, -1\}$.

Solution:

1. For $m = 0$, we should begin by finding the product sequence $s[n] = x[n]y[n]$. Doing this we get the following sequence:

   $$s[n] = \{\ldots, 0, 0, 2, 6, 24, 1, -9, 0, 0, \ldots\}$$

   and so from the sum in our crosscorrelation function we arrive at the answer of

   $$R_{xy}(0) = 22$$

2. For $m = 3$, we will approach it the same was we did above; however, we will now shift $y[n]$ to the right. Then we can find the

product sequence $s[n] = x[n]y[n-3]$, which yields

$$s[n] = \{\ldots, 0, 0, 0, 0, 0, 1, -6, 0, 0, \ldots\}$$

and from the crosscorrelation function we arrive at the answer of

$$R_{xy}(3) = -6$$

3. For $m = -1$, we will again take the same approach; however, we will now shift $y[n]$ to the left. Then we can find the product sequence $s[n] = x[n]y[n+1]$, which yields

$$s[n] = \{\ldots, 0, 0, -4, -12, 6, -3, 0, 0, 0, \ldots\}$$

and from the crosscorrelation function we arrive at the answer of

$$R_{xy}(-1) = -13$$

Cauchy-Schwarz Inequality
This module provides both statement and proof of the Cauchy-Schwarz inequality and discusses its practical implications with regard to the matched filter detector.

Introduction

Any treatment of linear algebra as relates to signal processing would not be complete without a discussion of the Cauchy-Schwarz ineqaulity, a relation that enables a wide array of signal procesing applications related to pattern matching through a method called the matched filter. Recall that in standard Euclidean space, the angle θ between two vectors x, y is given by
Equation:

$$\cos\left(\theta\right) = \frac{\langle x, y \rangle}{||x||||y||}.$$

Since $\cos\left(\theta\right) \leq 1$, it follows that
Equation:

$$\left|\langle x, y \rangle\right|^2 \leq \langle x, x \rangle \langle y, y \rangle.$$

Furthermore, equality holds if and only if $\cos\left(\theta\right) = 0$, implying that
Equation:

$$\left|\langle x, y \rangle\right|^2 = \langle x, x \rangle \langle y, y \rangle$$

if and only if $y = ax$ for some real a. This relation can be extended to all inner product spaces over a real or complex field and is known as the Cauchy-Schwarz inequality, which is of great importance to the study of signals.

The Cauchy-Schwarz Inequality

Statement of the Cauchy-Schwarz Inequality

The general statement of the Cauchy-Schwarz inequality mirrors the intuition for standard Euclidean space. Let V be an inner product space over the field of

complex numbers $\mathbb{C}$ with inner product $\langle \cdot, \cdot \rangle$. For every pair of vectors $x, y \in V$ the inequality
Equation:

$$|\langle x, y \rangle|^2 \leq \langle x, x \rangle \langle y, y \rangle$$

holds. Furthermore, the equality
Equation:

$$|\langle x, y \rangle|^2 = \langle x, x \rangle \langle y, y \rangle$$

holds if and only if $y = ax$ for some $a \in \mathbb{C}$. That is, equality holds if and only if x and y are linearly dependent.

Proof of the Cauchy-Schwarz Inequality

Let V be a vector space over the real or complex field F, and let $x, y \in V$ be given. In order to prove the Cauchy-Schwarz inequality, it will first be proven that $|\langle x, y \rangle|^2 = \langle x, x \rangle \langle y, y \rangle$ if $y = ax$ for some $a \in F$. It will then be shown that $|\langle x, y \rangle|^2 < \langle x, x \rangle \langle y, y \rangle$ if $y \neq ax$ for all $a \in F$.

Consider the case in which $y = ax$ for some $a \in F$. From the properties of inner products, it is clear that
Equation:

$$\begin{aligned} |\langle x, y \rangle|^2 &= |\langle x, ax \rangle|^2 \\ &= |a \langle x, x \rangle|^2. \end{aligned}$$

Hence, it follows that
Equation:

$$\begin{aligned} |\langle x, y \rangle|^2 &= |a^{\ 2}| \langle x, x \rangle|^2 \\ &= |a|^2 \langle x, x \rangle^2. \end{aligned}$$

Similarly, it is clear that
Equation:

$$\begin{aligned}\langle x,x\rangle\langle y,y\rangle &= \langle x,x\rangle\langle ax,ax\rangle \\ &= \langle x,x\rangle aa\,\langle x,x\rangle \\ &= |a|^2\langle x,x\rangle^2.\end{aligned}$$

Thus, it is proven that $|\langle x,y\rangle|^2 = \langle x,x\rangle\,\langle y,y\rangle$ if $x = ay$ for some $a \in F$.

Next, consider the case in which $y \neq ax$ for all $a \in F$, which implies that $y \neq 0$ so $\langle y,y\rangle \neq 0$. Thus, it follows by the properties of inner products that, for all $a \in F$, $\langle x-ay, x-ay\rangle > 0$. This can be expanded using the properties of inner products to the expression
Equation:

$$\begin{aligned}\langle x-ay,x-ay\rangle &= \langle x,x-ay\rangle - a\langle y,x-ay\rangle \\ &= \langle x,x\rangle - a\,\langle x,y\rangle - a\,\langle y,x\rangle + |a|^2\,\langle y,y\rangle\end{aligned}$$

Choosing $a = \frac{\langle x,y\rangle}{\langle y,y\rangle}$,
Equation:

$$\begin{aligned}\langle x-ay,x-ay\rangle &= \langle x,x\rangle - \frac{\langle y,x\rangle}{\langle y,y\rangle}\langle x,y\rangle - \frac{\langle x,y\rangle}{\langle y,y\rangle}\langle y,x\rangle + \frac{\langle x,y\rangle\langle y,x\rangle}{\langle y,y\rangle^2}\langle y,y\rangle \\ &= \langle x,x\rangle - \frac{\langle x,y\rangle\langle y,x\rangle}{\langle y,y\rangle}\end{aligned}$$

Hence, it follows that $\langle x,x\rangle - \frac{\langle x,y\rangle\langle y,x\rangle}{\langle y,y\rangle} > 0$. Consequently,
$\langle x,x\rangle\,\langle y,y\rangle - \langle x,y\rangle\langle x,y\rangle > 0$. Thus, it can be concluded that $|\langle x,y\rangle|^2 < \langle x,x\rangle\,\langle y,y\rangle$ if $y \neq ax$ for all $a \in F$.

Therefore, the inequality
Equation:

$$|\langle x,y\rangle|^2 \leq \langle x,x\rangle\,\langle y,y\rangle$$

holds for all $x, y \in V$, and equality
Equation:

$$|\langle x, y\rangle|^2 = \langle x, x\rangle \langle y, y\rangle$$

holds if and only if $y = ax$ for some $a \in F$.

Additional Mathematical Implications

Consider the maximization of $\left\langle \frac{x}{||x||}, \frac{y}{||y||} \right\rangle$ where the norm $||\cdot|| = \langle \cdot, \cdot \rangle$ is induced by the inner product. By the Cauchy-Schwarz inequality, we know that $\left\langle \frac{x}{||x||}, \frac{y}{||y||} \right\rangle^2 \leq 1$ and that $\left\langle \frac{x}{||x||}, \frac{y}{||y||} \right\rangle^2 = 1$ if and only if $\frac{y}{||y||} = a\frac{x}{||x||}$ for some $a \in \mathbb{C}$. Hence, $\left\langle \frac{x}{||x||}, \frac{y}{||y||} \right\rangle$ attains a maximum where $\frac{y}{||y||} = a\frac{x}{||x||}$ for some $a \in \mathbb{C}$. Thus, collecting the scalar variables, $\left\langle \frac{x}{||x||}, \frac{y}{||y||} \right\rangle$ attains a maximum where $y = ax$. This result will be particulaly useful in developing the matched filter detector techniques.

Matched Filter Detector

Background Concepts

A great many applications in signal processing, image processing, and beyond involve determining the presence and location of a target signal within some other signal. A radar system, for example, searches for copies of a transmitted radar pulse in order to determine the presence of and distance to reflective objects such as building or aircraft. A communication system searches for copies of waveforms representing digital 0s and 1s in order to receive a message.

As has already been shown, the expression $\left\langle \frac{x}{||x||}, \frac{y}{||y||} \right\rangle$ attains its upper bound, which is 1, when $y = ax$ for some scalar a in a real or complex field. The lower bound, which is 0, is attained when x and y are orthogonal. In informal intuition, this means that the expression is maximized when the vectors x and y have the same shape or pattern and minimized when x and y are very different. A pair of

vectors with similar but unequal shapes or patterns will produce relatively large value of the expression less than 1, and a pair of vectors with very different but not orthogonal shapes or patterns will produce relatively small values of the expression greater than 0. Thus, the above expression carries with it a notion of the degree to which two signals are “alike”, the magnitude of the normalized correlation between the signals in the case of the standard inner products.

This concept can be extremely useful. For instance consider a situation in which we wish to determine which signal, if any, from a set X of signals most resembles a particular signal y. In order to accomplish this, we might evaluate the above expression for every signal $x \in X$, choosing the one that results in maxima provided that those maxima are above some threshold of “likeness”. This is the idea behind the matched filter detector, which compares a set of signals against a target signal using the above expression in order to determine which among them are most like the target signal. For a detailed treatment of the applications of the matched filter detector see the liked module.

Signal Comparison

The simplest variant of the matched filter detector scheme would be to find the member signal in a set X of signals that most closely matches a target signal y. Thus, for every $x \in X$ we wish to evaluate
Equation:

$$m\left(x,y\right) = \left\langle \frac{x}{||x||}, \frac{y}{||y||} \right\rangle$$

in order to compare every member of X to the target signal y. Since the member of X which most closely matches the target signal y is desired, ultimately we wish to evaluate
Equation:

$$x_m = \text{argmax}_{x \in X} \left\langle \frac{x}{||x||}, \frac{y}{||y||} \right\rangle .$$

Note that the target signal does not technically need to be normalized to produce a maximum, but gives the desirable property that $m(x,y)$ is bounded to $[0,1]$.

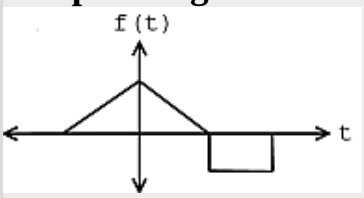
The element $x_m \in X$ that produces the maximum value of $m(x, y)$ is not necessarily unique, so there may be more than one matching signal in X. Additionally, the signal $x_m \in X$ producing the maximum value of $m(x, y)$ may not produce a very large value of $m(x, y)$ and thus not be very much like the target signal y. Hence, another matched filter scheme might identify the argument producing the maximum but only above a certain threshold, returning no matching signals in X if the maximum is below the threshold. There also may be a signal $x \in X$ that produces a large value of $m(x, y)$ and thus has a high degree of "likeness" to y but does not produce the maximum value of $m(x, y)$. Thus, yet another matched filter scheme might identify all signals in X producing local maxima that are above a certain threshold.

Example:
For example, consider the target signal given in [link] and the set of two signals given in [link]. By inspection, it is clear that the signal g_2 is most like the target signal f. However, to make that conclusion mathematically, we use the matched filter detector with the L_2 inner product. If we were to actually make the necessary computations, we would first normalize each signal and then compute the necessary inner products in order to compare the signals in X with the target signal f. We would notice that the absolute value of the inner product for g_2 with f when normalized is greater than the absolute value of the inner product of g_1 with f when normalized, mathematically stated as

Equation:

$$g_2 = \text{argmax}_{x \in \{g_1, g_2\}} \left\langle \frac{x}{||x||}, \frac{f}{||f||} \right\rangle .$$

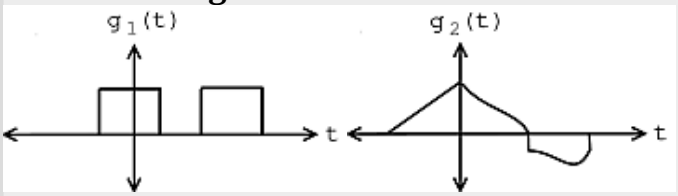
Template Signal



We wish to find a match for this target signal in the

set of signals
below.

Candidate Signals



We wish to find a match for the above
target signal in this set of signals.

Pattern Detection

A somewhat more involved matched filter detector scheme would involve attempting to match a target time limited signal $y = f$ to a set of of time shifted and windowed versions of a single signal $X = \{wS_t g | t \in \mathbb{R}\}$ indexed by $\mathbb{R}$. The windowing funtion is given by $w(t) = u(t - t_1) - u(t - t_2)$ where $[t_1, t_2]$ is the interval to which f is time limited. This scheme could be used to find portions of g that have the same shape as f. If the absolute value of the inner product of the normalized versions of f and $wS_t g$ is large, which is the absolute value of the normalized correlation for standard inner products, then g has a high degree of "likeness" to f on the interval to which f is time limited but left shifted by t. Of course, if f is not time limited, it means that the entire signal has a high degree of "likeness" to f left shifted by t.

Thus, in order to determine the most likely locations of a signal with the same shape as the target signal f in a signal g we wish to compute
Equation:

$$t_m = \text{argmax}_{t \in \mathbb{R}} \left\langle \frac{f}{||f||}, \frac{wS_t g}{||wS_t g||} \right\rangle$$

to provide the desired shift. Assuming the inner product space examined is $L_2(\mathbb{R}$ (similar results hold for $L_2\left(\mathbb{R}\left[a,b\right)\right)$, $l_2\left(\mathbb{Z}\right)$, and $l_2\left(\mathbb{Z}\left[a,b\right)\right)$), this produces
Equation:

$$t_m = \text{argmax}_{t\in\mathbb{R}} \ \frac{1}{||f||||wS_tg\,||}\int_{-\infty}^{\infty} f\left(\tau\right)w\left(\tau\right)g(\tau-t)d\tau\ .$$

Since f and w are time limited to the same interval
Equation:

$$t_m = \text{argmax}_{t\in\mathbb{R}} \ \frac{1}{||f||||wS_tg\,||}\int_{t_1}^{t_2} f\left(\tau\right)g(\tau-t)d\tau\ .$$

Making the subsitution $h\left(t\right)=g(-t)$,
Equation:

$$t_m = \text{argmax}_{t\in\mathbb{R}} \ \frac{1}{||f||||wS_tg\,||}\int_{t_1}^{t_2} f\left(\tau\right)h\left(t-\tau\right)d\tau\ .$$

Noting that this expression contains a convolution operation
Equation:

$$t_m = \text{argmax}_{t\in\mathbb{R}} \ \frac{(f*h)(t)}{||f||||wS_tg\,||}\ .$$

where h is the conjugate of the time reversed version of g defined by $h\left(t\right)=g(-t)$.

In the special case in which the target signal f is not time limited, w has unit value on the entire real line. Thus, the norm can be evaluated as $||wS_tg||=||S_tg||=||g||=||h||$. Therefore, the function reduces to $t_m = \text{argmax}_{t\in\mathbb{R}} \ \frac{(f*h)(t)}{||f||||h||}$ where $h\left(t\right)=g(-t)$. The function $f\star g=\frac{(f*h)(t)}{||f||||h||}$ is known as the normalized cross-correlation of f and g.

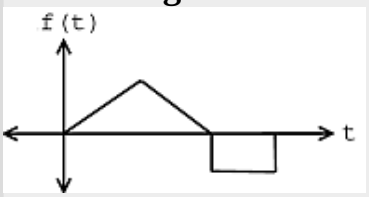
Hence, this matched filter scheme can be implemented as a convolution. Therefore, it may be expedient to implement it in the frequency domain. Similar results hold for the $L_2\left(\mathbb{R}\left[a,b\right)\right)$, $l_2\left(\mathbb{Z}\right)$, and $l_2\left(\mathbb{Z}\left[a,b\right]\right)$ spaces. It is especially useful to implement the $l_2\left(\mathbb{Z}\left[a,b\right]\right)$ cases in the frequency domain as the power of the Fast Fourier Transform algorithm can be leveraged to quickly perform the computations in a computer program. In the $L_2\left(\mathbb{R}\left[a,b\right)\right)$ and $l_2\left(\mathbb{Z}\left[a,b\right]\right)$ cases, care must be taken to zero pad the signal if wrap-around effects are not desired. Similar results also hold for spaces on higher dimensional intervals with the same inner products.

Of course, there is not necessarily exactly one instance of a target signal in a given signal. There could be one instance, more than one instance, or no instance of a target signal. Therefore, it is often more practical to identify all shifts corresponding to local maxima that are above a certain threshold.

Example:
The signal in [link] contains an instance of the template signal seen in [link] beginning at time $t = s_1$ as shown by the plot in [link]. Therefore,
Equation:

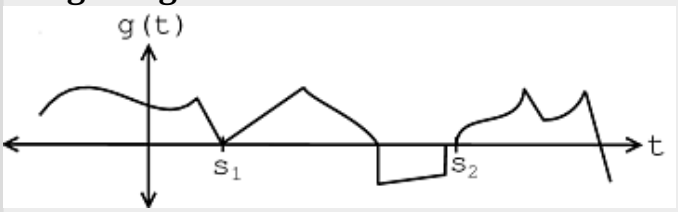
$$s_1 = \text{argmax}_{t\in\mathbb{R}} \left\langle \frac{f}{||f||}, \frac{wS_tg}{||wS_tg||} \right\rangle .$$

Pattern Signal



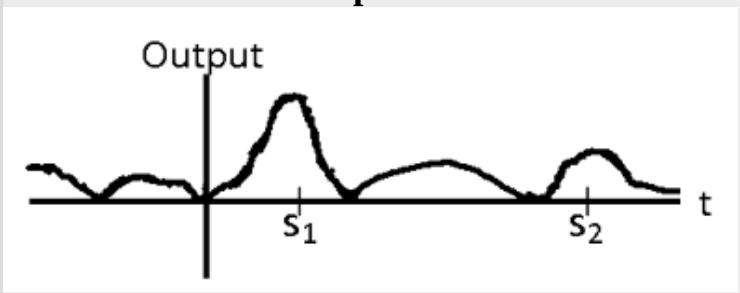
This function shows tha pattern we are looking for in the signal below, which occurs at time $t = s_1$.

Longer Signal



This signal contains an instance of the above signal starting at time $t = s_1$.

Absolute Value of Output



This signal shows a sketch of the absolute value of the matched filter output for the interval shown. Note that this was just an "eyeball approximation" sketch. Observe the pronounced peak at time $t = s_1$.

Cauchy-Schwarz Inequality Video Lecture

Proof of the Cauchy-Schwarz Inequality

[missing_resource: http://www.youtube.com/v/r2PogGDl8_U&hl=en_US&fs=1&rel=0]

Video lecture on the proof of the Cauchy-Schwarz inequality from Khan Academy. Only part of the theorem is proven.

Cauchy-Schwarz Inequality Summary

As can be seen, the Cauchy-Schwarz inequality is a property of inner product spaces over real or complex fields that is of particular importance to the study of signals. Specifically, the implication that the absolute value of an inner product is maximized over normal vectors when the two arguments are linearly dependent is key to the justification of the matched filter detector. Thus, it enables the use of matched filters for such pattern matching applications as image detection, communications demodulation, and radar signal analysis.

The Q-function

The Q-function is a convenient way to express right-tail probabilities for normal (Gaussian) random variables. For $x \in \mathbb{R}$, $Q(x)$ is defined as the probability that a standard normal random variable (zero mean, unit variance) exceeds x:

$$Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{-\frac{t^2}{2}} \, \mathrm{d}\, t$$

Q is a mapping from $\mathbb{R}$ to $\{0, 1\}$. One may also define $Q(-\infty) = 1$ and $Q(\infty) = 0$.

Note: Some authors define the Q-function in different ways. One alternative is to define $Q(x) = F(x) - F(0)$. This definition is discussed at MathWorld.

[missing_resource: .png]

$Q(x)$ is represented by the shaded region.

Since $Q(x)$ is monotonically decreasing, it has a well-defined inverse $Q^{-1} : \{0, 1\} \to \mathbb{R}$.
[missing_resource: .png]

A plot of $Q(x)$

If $F(x)$ denotes the cumulative distribution function of a standard normal, then clearly $Q(x) = 1 - F(x)$. For this reason Q is also called the **complementary cumulative distribution function**. The Q-function is useful because the tail probability cannot be evaluated symbolically, and so $Q(x)$ offers a concise notation for this integral. It is similar to the gamma and beta functions in this respect.

Arbitrary Mean and Variance

The Q-function is also useful for expressing right-tail probabilities of nonstandard normal variates. If

$$X \sim \mathcal{N}\left(\mu, \sigma^2\right)$$

then

$$\frac{X - \mu}{\sigma} \sim \mathcal{N}(0, 1)$$

To express $\Pr[X > \gamma]$ in terms of Q, where $\gamma \in \mathbb{R}$, define $\eta = \frac{\gamma - \mu}{\sigma}$. Then
Equation:

$$\begin{aligned} \Pr[X > \gamma] &= \Pr[X > \eta\sigma + \mu] \\ &= \Pr\left[\tfrac{X-\mu}{\sigma} > \eta\right] \\ &= Q(\eta) \\ &= Q\left(\tfrac{\gamma - \mu}{\sigma}\right) \end{aligned}$$

Relation to Erf and Erfinv

The erf function is defined as

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} \, \mathrm{d}\, t$$

Both erf and its inverse, erfinv, are built into many common mathematical software packages such as Mathematica and Matlab. Therefore, they can be used to numerically evaluate Q and Q^{-1}. By a change of variables, we have

$$Q(x) = \frac{1}{2} \quad 1 - \operatorname{erf} \quad \frac{x}{\sqrt{2}}$$

and

$$Q^{-1}(y) = \sqrt{2}\operatorname{erfinv}\left(1 - 2y\right)$$

Approximations

One approximation that is sometimes useful for x away from zero is

$$Q(x) \simeq \frac{1}{x\sqrt{2\pi}} e^{\frac{-1}{2}x^2}$$

Introduction to Detection Theory

Introduction

The intent of **detection theory** is to provide rational (instead of arbitrary) techniques for determining which of several conceptions--models--of data generation and measurement is most "consistent" with a given set of data. In digital communication, the received signal must be processed to determine whether it represents a binary "0" or "1"; in radar or sonar, the presence or absence of a target must be determined from measurements of propagating fields; in seismic problems, the presence of oil deposits must be inferred from measurements of sound propagation in the earth. Using detection theory, we will derive signal processing algorithms which will give good answers to questions such as these when the information-bearing signals are corrupted by superfluous signals (noise).

The detection theory's foundation rests on statistical hypothesis testing (Cramér, 1946, Chapter 35; Lehman, 1986; Poor, 1988, Chapter 2; van Trees, 1968, pp 19-52). Given a probabilistic model (an event space Ω and the associated probabilistic structures), a random vector $\boldsymbol{r}$ expressing the observed data, and a listing of the probabilistic models--the **hypotheses**--which may have generated $\boldsymbol{r}$, we want a systematic, optimal method of determining which model corresponds to the data. In the simple case where only two models-- $\mathcal{M}_0$ and $\mathcal{M}_1$--are possible, we ask, for each set of observations, what is the "best" method of deciding whether $\mathcal{M}_0$ or $\mathcal{M}_1$ was true? We have many ways of mathematically stating what "best" means: we shall initially choose the average cost of each decision as our criterion for correctness. This seemingly arbitrary choice of criterion will be shown later **not** to impose rigid constraints on the algorithms that solve the hypothesis testing problem. Over a variety of reasonable criteria, one central solution to evaluating which model describes observations--the likelihood ratio test--will persistently emerge; this result will form the basis of **all** detection algorithms.

Detection problems become more elaborate and complicated when models become vague. Models are characterized by probability distributions, and these distributions suffice in the likelihood ratio test. Vagueness does not

refer to this stochastic framework; rather, it refers to uncertainties in the probability distribution itself. The distribution may depend on unknown parameters, like noise power level. The distribution most certainly depends on signal structure; suppose that is partially or completely unknown? The most difficult (and interesting) problems emerge when uncertainties arise in the probability distributions themselves. For example, suppose the only model information we have is through data; how would an optimal detector be derived then?

Along the way we will discover that a general geometric picture of detection emerges: Ease of a detection problem depends on how "far apart" the models are from each other. This geometric framework turns out to be elaborate, but underlies modern detection theory and forms links to information theory.

Discrete-Time Detection Theory

Introduction

Detection theory applies optimal model evaluation to signals (Helstrom, Poor, van Trees). Usually, we measure a signal in the presence of additive noise over some finite number of samples. Each observed datum is of the form $s(l) + n(l)$, where $s(l)$ denotes the l^{th} signal value and $n(l)$ the l^{th} noise value. In this and in succeeding sections of this chapter, we focus the general methods of evaluating models.

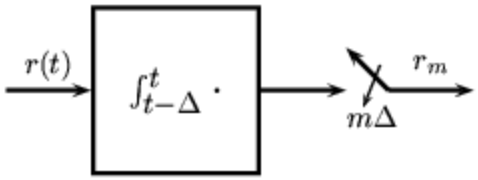
Detection of Signals in Gaussian Noise

For the moment, we assume we know the joint distribution of the noise values. In most cases, the various models for the form of the observations - the hypothesis - do not differ because of noise characteristics. Rather, the signal component determines model variations and the noise is statistically independent of the signal; such is the specificity of detection problems in contrast to the generality of model evaluation. For example, we may want to determine whether a signal characteristic of a particular ship is present in a sonar array's output (the signal is known) or whether no ship is present (zero-valued signal).

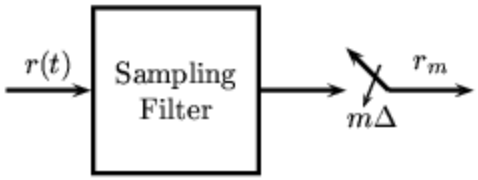
To apply optimal hypothesis testing procedures previously derived, we first obtain a finite number L of observations $r(l), l \in \{0, \ldots, L-1\}$. These observations are usually obtained from continuous-time observations in one of two ways. Two commonly used methods for passing from continuous-time to discrete-time are known: **integrate-and-dump** and **sampling**. These techniques are illustrated in [link].

The two most common methods of converting continuous-time observations into discrete-time ones are shown. In the upper panel, the integrate-and-dump method is shown: the input is integrated over an

interval of duration Δ and the result sampled. In the lower panel, the sampling method merely samples the input every Δ seconds.



Integrate-and-Dump



Sampling

Integrate-and-Dump

In this procedure, no attention is paid to the bandwidth of the noise in selecting the sampling rate. Instead, the sampling interval Δ is selected according to the characteristics of the signal set. Because of the finite duration of the integrator, successive samples are statistically independent when the noise bandwidth exceeds $\frac{1}{\Delta}$ Consequently, the sampling rate can be varied to some extent while retaining this desirable analytic property.

Sampling

Traditional engineering considerations governed the selection of the sampling filter and the sampling rate. As in the integrate-and-dump procedure, the sampling rate is chosen according to signal properties. Presumably, changes in sampling rate would force changes in the filter. As we shall see, this linkage has dramatic implications on performance.

With either method, the continuous-time detection problem of selecting between models (a binary selection is used here for simplicity)

$$\mathscr{M}_0 : r(t) = s^0(t) + n(t) \;\; 0 \leq t < T$$

$$\mathscr{M}_1 : r(t) = s^1(t) + n(t) \;\; 0 \leq t < T$$

where $\left\{s^i(t)\right\}$ denotes the known signal set and $n(t)$ denotes additive noise modeled as a stationary stochastic process [footnote] is converted into the discrete-time detection problem

$$\mathscr{M}_0 : r_l = s_l^0 + n_l \;\; 0 \leq l < L$$

$$\mathscr{M}_1 : r_l = s_l^1 + n_l \;\; 0 \leq l < L$$

where the sampling interval is **always** taken to divide the observation interval $T : L = \frac{T}{\Delta}$. We form the discrete-time observations into a vector: $\boldsymbol{r} = (r(0) \ldots r(L-1))^T$. The binary detection problem is to distinguish between two possible signals present in the noisy output waveform.

$$\mathscr{M}_0 : \boldsymbol{r} = s_0 + \boldsymbol{n}$$

$$\mathscr{M}_0 : \boldsymbol{r} = s_1 + \boldsymbol{n}$$

To apply our model evaluation results, we need the probability density of $\boldsymbol{r}$ under each model. As the only probabilistic component of the observations is the noise, the required density for the detection problem is given by

$$\mathrm{p}_{\boldsymbol{r}\,|\,\mathscr{M}_i}\,(\boldsymbol{r}) = \mathrm{p}_{\boldsymbol{n}}\,(\boldsymbol{r} - s_i)$$

and the corresponding likelihood ratio by

$$\Lambda(\boldsymbol{r}) = \frac{\mathrm{p}_{\boldsymbol{n}}\left(\boldsymbol{r} - \boldsymbol{s}_1\right)}{\mathrm{p}_{\boldsymbol{n}}\left(\boldsymbol{r} - \boldsymbol{s}_0\right)}$$

Much of detection theory revolves about interpreting this likelihood ratio and deriving the detection threshold (either $\mathrm{threshold}$ or γ).
We are **not** assuming the amplitude distribution of the noise to be Gaussian.

Hypothesis Testing

Suppose you measure a collection of scalars $x_1, \ldots, x_N$. You believe the data is distributed in one of two ways. Your first model, call it H_0, postulates the data to be governed by the density $f_0(x)$ (some fixed density). Your second model, H_1, postulates a different density $f_1(x)$. These models, termed **hypotheses**, are denoted as follows:

$$H_0 : x_n \sim f_0(x), n = 1 \ldots N$$

$$H_1 : x_n \sim f_1(x), n = 1 \ldots N$$

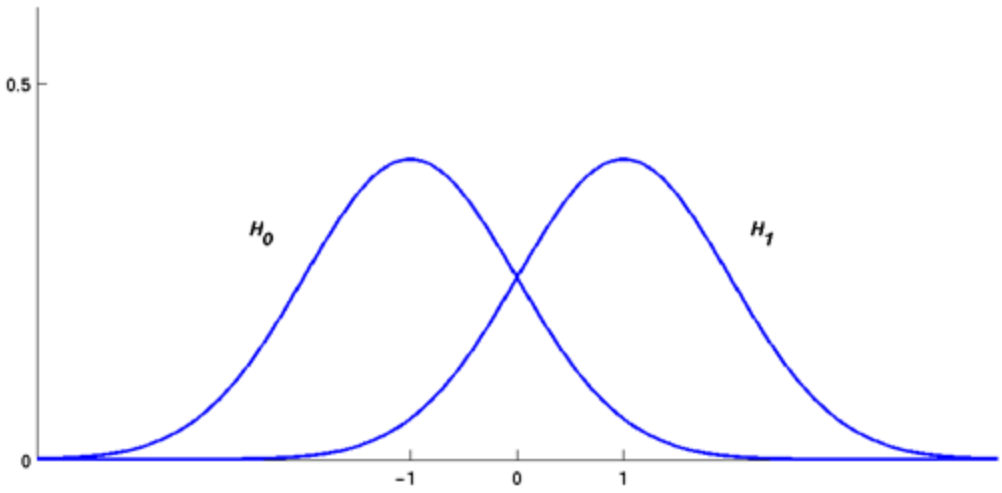
A **hypothesis test** is a rule that, given a measurement $\boldsymbol{x}$, makes a decision as to which hypothesis best "explains" the data.

Example:
Suppose you are confident that your data is normally distributed with variance 1, but you are uncertain about the sign of the mean. You might postulate

$$H_0 : x_n \sim \mathcal{N}(-1, 1)$$

$$H_1 : x_n \sim \mathcal{N}(1, 1)$$

These densities are depicted in [link].



Assuming each hypothesis is a priori equally likely, an intuitively appealing hypothesis test is to compute the sample mean $\bar{x} = \frac{1}{N} \sum_{n=1}^{N} x_n$, and choose H_0 if $\bar{x} \leq 0$, and H_1 if $\bar{x} > 0$. As we will see later, this test is in fact optimal under certain assumptions.

Generalizations and Nomenclature

The concepts introduced above can be extended in several ways. In what follows we provide more rigorous definitions, describe different kinds of hypothesis testing, and introduce terminology.

Data

In the most general setup, the observation is a collection $x_1, \ldots, x_N$ of random vectors. A common assumption, which facilitates analysis, is that the data are independent and identically distributed (IID). The random vectors may be continuous, discrete, or in some cases mixed. It is generally assumed that all of the data is available at once, although for some applications, such as Sequential Hypothesis Testing, the data is a never ending stream.

Binary Versus M-ary Tests

When there are two competing hypotheses, we refer to a **binary** hypothesis test. When the number of hypotheses is $M \geq 2$, we refer to an **M-ary** hypothesis test. Clearly, binary is a special case of M-ary, but binary tests are accorded a special status for certain reasons. These include their simplicity, their prevalence in applications, and theoretical results that do not carry over to the M-ary case.

Example:

Phase-Shift Keying

Suppose we wish to transmit a binary string of length r over a noisy communication channel. We assign each of the $M = 2^r$ possible bit sequences to a signal s^k, $k = \{1, \ldots, M\}$ where

$$s_n^k = \cos\left(2\pi f_0 n + \frac{2\pi\,(k-1)}{M}\right)$$

This symboling scheme is known as **phase-shift keying** (PSK). After transmitting a signal across the noisy channel, the receiver faces an M-ary hypothesis testing problem:

$$H_0 : \boldsymbol{x} = s^1 + \boldsymbol{w}$$

$$\vdots$$

$$H_M : \boldsymbol{x} = s^M + \boldsymbol{w}$$

where $\boldsymbol{w} \sim \mathcal{N}\left(\mathbf{0}, \sigma^2 I\right)$.

In many binary hypothesis tests, one hypothesis represents the absence of a ceratin feature. In such cases, the hypothesis is usually labelled H_0 and called the **null** hypothesis. The other hypothesis is labelled H_1 and called the **alternative** hypothesis.

Example:

Waveform Detection

Consider the problem of detecting a known signal $\boldsymbol{s} = \left(s_1 \ldots s_N\right)^T$ in additive white Gaussian noise (AWGN). This scenario is common in sonar and radar systems. Denoting the data as $\boldsymbol{x} = \left(x_1 \ldots x_N\right)^T$, our hypothesis testing problem is

$$H_0 : \boldsymbol{x} = \boldsymbol{w}$$

$$H_1 : \boldsymbol{x} = \boldsymbol{s} + \boldsymbol{w}$$

where $\boldsymbol{w} \sim \mathcal{N}\left(\mathbf{0}, \sigma^2 I\right)$. H_0 is the null hypothesis, corresponding to the absence of a signal.

Tests and Decision Regions

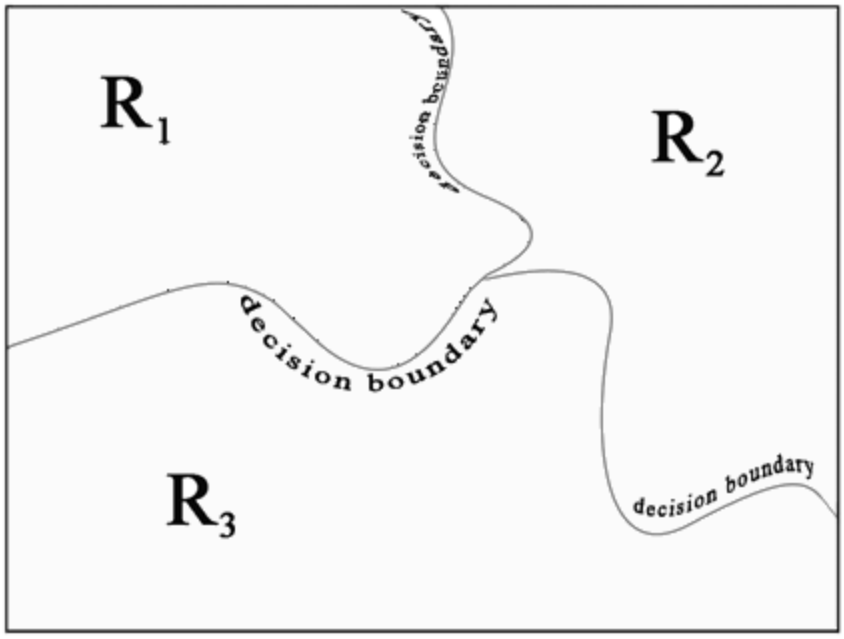
Consider the general hypothesis testing problem where we have N d-dimensional observations $x_1, \ldots, x_N$ and M hypotheses. If the data are real-valued, for example, then a hypothesis test is a mapping

$$\varphi : \left(\mathbb{R}^d\right)^N \to \{1, \ldots, M\}$$

For every possible realization of the input, the test outputs a hypothesis. The test φ partitions the input space into a disjoint collection $R_1, \ldots, R_M$, where

$$R_k = \{(x_1, \ldots, x_N) | \varphi(x_1, \ldots, x_N) = k\}$$

The sets R_k are called **decision regions**. The boundary between two decision regions is a **decision boundary**. [link] depicts these concepts when $d = 2$, $N = 1$, and $M = 3$.



Simple Versus Composite Hypotheses

If the distribution of the data under a certain hypothesis is fully known, we call it a **simple** hypothesis. All of the hypotheses in the examples above are simple. In many cases, however, we only know the distribution up to certain unknown parameters. For example, in a Gaussian noise model we may not know the variance of the noise. In this case, a hypothesis is said to be **composite**.

Example:
Consider the problem of detecting the signal

$$s_n = \cos(2\pi f_0 \, (n-k)) \forall n : (n = \{1, \ldots, N\})$$

where k is an unknown delay parameter. Then

$$H_0 : \boldsymbol{x} = \boldsymbol{w}$$

$$H_1 : \boldsymbol{x} = \boldsymbol{s} + \boldsymbol{w}$$

is a binary test of a simple hypothesis (H_0) versus a composite alternative. Here we are assuming $w_n \sim \mathcal{N}\left(0, \sigma^2\right)$, with σ^2 known.

Often a test involving a composite hypothesis has the form

$$H_0 : \boldsymbol{\theta} = \theta_0$$

$$H_1 : \boldsymbol{\theta} \neq \theta_0$$

where θ_0 is fixed. Such problems are called **two-sided** because the composite alternative "lies on both sides of H_0." When $\boldsymbol{\theta}$ is a scalar, the test

$$H_0 : \theta \leq \theta_0$$

$$H_1 : \theta > \theta_0$$

is called **one-sided**. Here, both hypotheses are composite.

Example:
Suppose a coin turns up heads with probability p. We want to assess whether the coin is fair ($p = \frac{1}{2}$). We toss the coin N times and record $x_1, \ldots, x_N$ ($x_n = 1$ means heads and $x_n = 0$ means tails). Then

$$H_0 : p = \frac{1}{2}$$

$$H_1 : p \neq \frac{1}{2}$$

is a binary test of a simple hypothesis (H_0) versus a composite alternative. This is also a two-sided test.

Errors and Probabilities

In binary hypothesis testing, assuming at least one of the two models does indeed correspond to reality, there are four possible scenarios:

- **Case 1** H_0 is true, and we declare H_0 to be true
- **Case 2** H_0 is true, but we declare H_1 to be true
- **Case 3** H_1 is true, and we declare H_1 to be true
- **Case 4** H_1 is true, but we declare H_0 to be true

In cases 2 and 4, errors occur. The names given to these errors depend on the area of application. In statistics, they are called **type I** and **type II errors** respectively, while in signal processing they are known as a **false alarm** or a **miss**.

Consider the general binary hypothesis testing problem

$$H_0 : \boldsymbol{x} \sim f_\theta(\boldsymbol{x}), \theta \in \Theta_0$$

$$H_1 : \boldsymbol{x} \sim f_\theta(\boldsymbol{x}), \theta \in \Theta_1$$

If H_0 is simple, that is, $\Theta_0 = \{\theta_0\}$, then the **size** (denoted α), also called the **false-alarm probability** (P_F), is defined to be

$$\alpha = P_F = \Pr[\theta_0, \text{declare} H_1]$$

When Θ_0 is composite, we define

$$\alpha = P_F = \sup_{\theta \in \Theta_0} (\Pr[\theta, \text{declare} H_1])$$

For $\theta \in \Theta_1$, the **power** (denoted β), or **detection probability** (P_D), is defined to be

$$\beta = P_D = \Pr[\theta, \text{declare} H_1]$$

The probability of a type II error, also called the **miss probability**, is

$$P_M = 1 - P_D$$

If H_1 is composite, then $\beta = \beta(\theta)$ is viewed as a function of θ.

Criteria in Hypothesis Testing

The design of a hypothesis test/detector often involves constructing the solution to an optimization problem. The optimality criteria used fall into two classes: Bayesian and frequent.

Representing the former approach is the Bayes Risk Criterion. Representing the latter is the Neyman-Pearson Criterion. These two approaches are developed at length in separate modules.

Statistics Versus Engineering Lingo

The following table, adapted from Kay, p.65, summarizes the different terminology for hypothesis testing from statistics and signal processing:

Statistics	**Signal Processing**
Hypothesis Test	Detector
Null Hypothesis	Noise Only Hypothesis
Alternate Hypothesis	Signal + Noise Hypothesis
Critical Region	Signal Present Decision Region
Type I Error	False Alarm
Type II Error	Miss
Size of Test (α)	Probability of False Alarm (P_F)
Power of Test (β)	Probability of Detection (P_D)

The Neyman-Pearson Criterion

In hypothesis testing, as in all other areas of statistical inference, there are two major schools of thought on designing good tests: Bayesian and frequentist (or classical). Consider the simple binary hypothesis testing problem

$$\mathscr{H}_0 : \boldsymbol{x} \sim f_0(\boldsymbol{x})$$

$$\mathscr{H}_1 : \boldsymbol{x} \sim f_1(\boldsymbol{x})$$

In the Bayesian setup, the prior probability $\pi_i = \Pr[\mathscr{H}_i]$ of each hypothesis occurring is assumed known. This approach to hypothesis testing is represented by the minimum Bayes risk criterion and the minimum probability of error criterion.

In some applications, however, it may not be reasonable to assign an a priori probability to a hypothesis. For example, what is the a priori probability of a supernova occurring in any particular region of the sky? What is the prior probability of being attacked by a ballistic missile? In such cases we need a decision rule that does not depend on making assumptions about the a priori probability of each hypothesis. Here the Neyman-Pearson criterion offers an alternative to the Bayesian framework.

The Neyman-Pearson criterion is stated in terms of certain probabilities associated with a particular hypothesis test. The relevant quantities are summarized in [link]. Depending on the setting, different terminology is used.

Statistics		**Signal Processing**		
Probability	**Name**	**Notation**	**Name**	**Notation**
$P_0(\text{declare}\mathscr{H}_1)$	size	α	false-alarm probability	P_F

Statistics		Signal Processing		
Probability	**Name**	**Notation**	**Name**	**Notation**
$P_1(\text{declare}\mathscr{H}_1)$	power	β	detection probability	P_D

Here $P_i(\text{declare}\mathscr{H}_j)$ dentoes the probability that we declare hypothesis $\mathscr{H}_j$ to be in effect when $\mathscr{H}_i$ is actually in effect. The probabilities $P_0(\text{declare}\mathscr{H}_0)$ and $P_1(\text{declare}\mathscr{H}_0)$ (sometimes called the **miss** probability), are equal to $1 - P_F$ and $1 - P_D$, respectively. Thus, P_F and P_D represent the two degrees of freedom in a binary hypothesis test. Note that P_F and P_D do not involve a priori probabilities of the hypotheses.

These two probabilities are related to each other through the decision regions. If R_1 is the decision region for $\mathscr{H}_1$, we have

$$P_F = \int f_0(\boldsymbol{x}) \, \mathrm{d}\, \boldsymbol{x}$$

$$P_D = \int f_1(\boldsymbol{x}) \, \mathrm{d}\, \boldsymbol{x}$$

The densities $f_i(\boldsymbol{x})$ are nonnegative, so as R_1 shrinks, both probabilities tend to zero. As R_1 expands, both tend to one. The ideal case, where $P_D = 1$ and $P_F = 0$, cannot occur unless the distributions do not overlap (i.e., $\int f_0(\boldsymbol{x}) f_1(\boldsymbol{x}) \, \mathrm{d}\, \boldsymbol{x} = 0$). Therefore, in order to increase P_D, we must also increase P_F. This represents the fundamental tradeoff in hypothesis testing and detection theory.

Example:
Consider the simple binary hypothesis test of a scalar measurement x:

$$\mathscr{H}_0 : x \sim \mathcal{N}(0, 1)$$

$$\mathscr{H}_1 : x \sim \mathcal{N}(1, 1)$$

Suppose we use a threshold test

$$x \underset{\mathscr{H}_0}{\overset{\mathscr{H}_1}{\gtrless}} \gamma$$

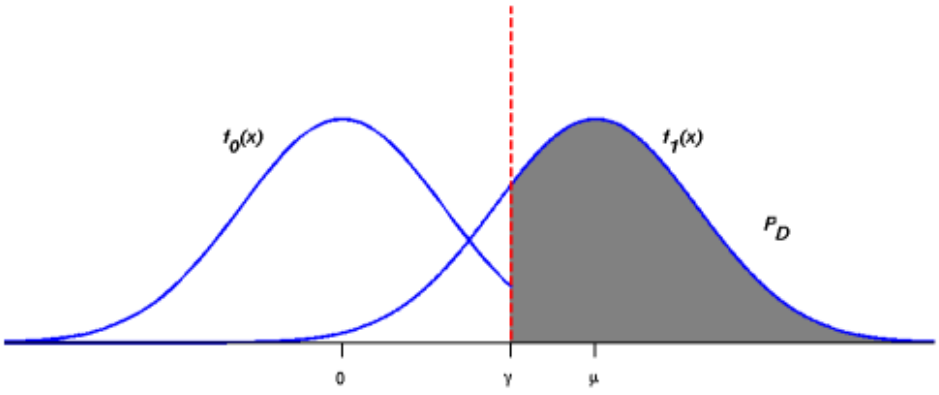
where $\gamma \in \mathbb{R}$ is a free parameter. Then the false alarm and detection probabilities are

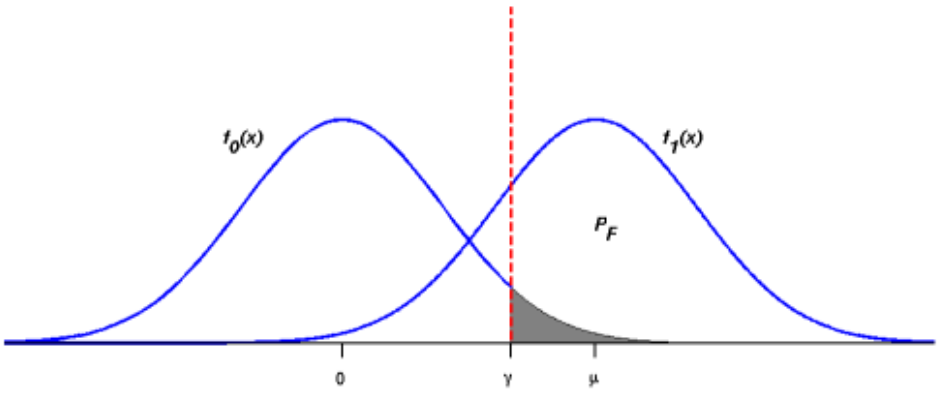
$$P_F = \int_\gamma^\infty \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} \, \mathrm{d}\, t = Q(\gamma)$$

$$P_D = \int_\gamma^\infty \frac{1}{\sqrt{2\pi}} e^{-\frac{(t-1)^2}{2}} \, \mathrm{d}\, t = Q(\gamma - 1)$$

where Q denotes the Q-function. These quantities are depicted in [link].

False alarm and detection values for a certain threshold.

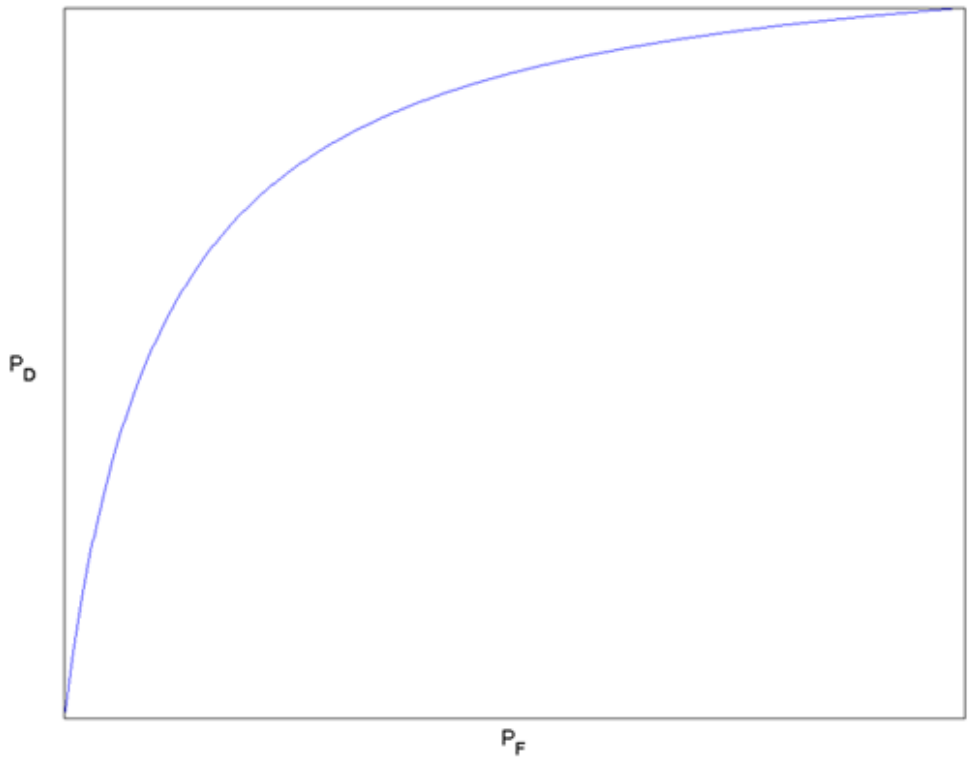




Since the Q-function is monotonicaly decreasing, it is evident that both P_D and P_F decay to zero as γ increases. There is also an explicit relationship

$$P_D = Q\left(Q^{-1}(P_F) - 1\right)$$

A common means of displaying this relationship is with a **receiver operating characteristic** (ROC) curve, which is nothing more than a plot of P_D versus P_F ([link]).



ROC curve for this example.

The Neyman-Pearson Lemma: A First Look

The Neyman-Pearson criterion says that we should construct our decision rule to have maximum probability of detection while not allowing the probability of false alarm to exceed a certain value α. In other words, the optimal detector according to the Neyman-Pearson criterion is the solution to the following constrainted optimization problem:

Neyman-Pearson Criterion

Equation:

$$\max \{P_D\}, \text{such that} P_F \leq \alpha$$

The maximization is over all decision rules (equivalently, over all decision regions R_0, R_1). Using different terminology, the Neyman-Pearson criterion selects the **most powerful test of size (not exceeding)** α.

Fortunately, the above optimization problem has an explicit solution. This is given by the celebrated **Neyman-Pearson lemma**, which we now state. To ease the exposition, our initial statement of this result only applies to continuous random variables, and places a technical condition on the densities. A more general statement is given later in the module.
Neyman-Pearson Lemma: initial statement

Consider the test

$$\mathscr{H}_0 : \boldsymbol{x} \sim f_0(\boldsymbol{x})$$

$$\mathscr{H}_1 : \boldsymbol{x} \sim f_1(\boldsymbol{x})$$

where $f_i(\boldsymbol{x})$ is a density. Define $\Lambda(\boldsymbol{x}) = \frac{f_1(\boldsymbol{x})}{f_0(\boldsymbol{x})}$, and assume that $\Lambda(\boldsymbol{x})$ satisfies the condition that for each $\gamma \in \mathbb{R}$, $\Lambda(\boldsymbol{x})$ takes on the value γ with probability zero under hypothesis $\mathscr{H}_0$. The solution to the optimization problem in [link] is given by

$$\Lambda(\boldsymbol{x}) = \frac{f_1(\boldsymbol{x})}{f_0(\boldsymbol{x})} \underset{\mathscr{H}_0}{\overset{\mathscr{H}_1}{\gtrless}} \eta$$

where η is such that

$$P_F = \int f_0(\boldsymbol{x}) \, \mathrm{d} \, \boldsymbol{x} = \alpha$$

If $\alpha = 0$, then $\eta = \infty$. The optimal test is unique up to a set of probability zero under $\mathscr{H}_0$ and $\mathscr{H}_1$.
The optimal decision rule is called the **likelihood ratio test**. $\Lambda(\boldsymbol{x})$ is the **likelihood ratio**, and η is a **threshold**. Observe that neither the likelihood

ratio nor the threshold depends on the a priori probabilities $\Pr[\mathscr{H}_i]$. they depend only on the conditional densities f_i and the size constraint α. The threshold can often be solved for as a function of α, as the next example shows.

Example:
Continuing with [link], suppose we wish to design a Neyman-Pearson decision rule with size constraint α. We have
Equation:

$$\begin{aligned} \Lambda(x) &= \frac{\frac{1}{\sqrt{2\pi}} e^{-\frac{(x-1)^2}{2}}}{\frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}} \\ &= e^{x-\frac{1}{2}} \end{aligned}$$

By taking the natural logarithm of both sides of the LRT and rarranging terms, the decision rule is not changed, and we obtain

$$x \underset{\mathscr{H}_0}{\overset{\mathscr{H}_1}{\gtrless}} \ln(\eta) + \frac{1}{2} \equiv \gamma$$

Thus, the optimal rule is in fact a thresholding rule like we considered in [link]. The false-alarm probability was seen to be

$$P_F = Q(\gamma)$$

Thus, we may express the value of γ required by the Neyman-Pearson lemma in terms of α:

$$\gamma = Q^{-1}(\alpha)$$

Sufficient Statistics and Monotonic Transformations

For hypothesis testing involving multiple or vector-valued data, direct evaluation of the size (P_F) and power (P_D) of a Neyman-Pearson decision rule would require integration over multi-dimensional, and potentially complicated decision regions. In many cases, however, this can be avoided by simplifying the LRT to a test of the form

$$\boldsymbol{t} \underset{\mathscr{H}_0}{\overset{\mathscr{H}_1}{\gtrless}} \gamma$$

where the test statistic $\boldsymbol{t} = T(\boldsymbol{x})$ is a sufficient statistic for the data. Such a simplified form is arrived at by modifying both sides of the LRT with montonically increasing transformations, and by algebraic simplifications. Since the modifications do not change the decision rule, we may calculate P_F and P_D in terms of the sufficient statistic. For example, the false-alarm probability may be written

Equation:

$$\begin{aligned} P_F &= \Pr[\text{declare} \mathscr{H}_1] \\ &= \int f_0(\boldsymbol{t}) \, \mathrm{d}\, \boldsymbol{t} \end{aligned}$$

where $f_0(\boldsymbol{t})$ denotes the density of $\boldsymbol{t}$ under $\mathscr{H}_0$. Since $\boldsymbol{t}$ is typically of lower dimension than $\boldsymbol{x}$, evaluation of P_F and P_D can be greatly simplified. The key is being able to reduce the LRT to a threshold test involving a sufficient statistic **for which we know the distribution**.

Example:

Common Variances, Uncommon Means

Let's design a Neyman-Pearson decision rule of size α for the problem

$$\mathscr{H}_0 : \boldsymbol{x} \sim \mathcal{N}\left(\mathbf{0}, \sigma^2 I\right)$$

$$\mathscr{H}_1 : \boldsymbol{x} \sim \mathcal{N}\left(\mu \mathbf{1}, \sigma^2 I\right)$$

where $\mu > 0$, $\sigma^2 > 0$ are known, $\mathbf{0} = (0\ldots0)^T$, $\mathbf{1} = (1\ldots1)^T$ are N-dimensional vectors, and I is the $N{\times}N$ identity matrix. The likelihood ratio is

Equation:

$$\begin{aligned}
\Lambda(\boldsymbol{x}) &= \frac{\prod_{\mathrm{n}=1}^{N} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x_n-\mu)^2}{2\sigma^2}}}{\prod_{\mathrm{n}=1}^{N} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x_n{}^2}{2\sigma^2}}} \\
&= \frac{e^{-\sum_{\mathrm{n}=1}^{N} \frac{(x_n-\mu)^2}{2\sigma^2}}}{e^{-\sum_{\mathrm{n}=1}^{N} \frac{x_n{}^2}{2\sigma^2}}} \\
&= e^{\frac{1}{2\sigma^2}\sum_{\mathrm{n}=1}^{N} 2x_n\mu - \mu^2} \\
&= e^{\frac{1}{\sigma^2}\left(-\frac{N\mu^2}{2} + \mu\sum_{\mathrm{n}=1}^{N} x_n\right)}
\end{aligned}$$

To simplify the test further we may apply the natural logarithm and rearrange terms to obtain

$$t \equiv \sum_{\mathrm{n}=1}^{N} x_n \underset{\mathscr{H}_0}{\overset{\mathscr{H}_1}{\gtrless}} \frac{\sigma^2}{\mu}\ln(\eta) + \frac{N\mu}{2} \equiv \gamma$$

Note: We have used the assumption $\mu > 0$. If $\mu < 0$, then division by μ is not a monotonically increasing operation, and the inequalities would be reversed.

The test statistic t is sufficient for the unknown mean. To set the threshold γ, we write the false-alarm probability (size) as

$$P_F = \Pr[t > \gamma] = \int f_0(t)\ \mathrm{d}\ t$$

To evaluate P_F, we need to know the density of t under $\mathscr{H}_0$. Fortunately, t is the sum of normal variates, so it is again

normally distributed. In particular, we have $t = \boldsymbol{A}\boldsymbol{x}$, where $\boldsymbol{A} = \mathbf{1}^T$, so

$$t \sim \mathcal{N}\left(\boldsymbol{A}\mathbf{0}, \boldsymbol{A}\left(\sigma^2 I\right)\boldsymbol{A}^T\right) = \mathcal{N}\left(0, N\sigma^2\right)$$

under $\mathscr{H}_0$. Therefore, we may write P_F in terms of the Q-function as

$$P_F = Q\left(\frac{\gamma}{\sqrt{N}\sigma}\right)$$

The threshold is thus determined by

$$\gamma = \sqrt{N}\sigma Q^{-1}(\alpha)$$

Under $\mathscr{H}_1$, we have

$$t \sim \mathcal{N}\left(\boldsymbol{A}\mathbf{1}, \boldsymbol{A}\left(\sigma^2 I\right)\boldsymbol{A}^T\right) = \mathcal{N}\left(N\mu, N\sigma^2\right)$$

and so the detection probability (power) is

$$P_D = \Pr[t > \gamma] = Q\left(\frac{\gamma - N\mu}{\sqrt{N}\sigma}\right)$$

Writing P_D as a function of P_F, the ROC curve is given by

$$P_D = Q\left(Q^{-1}(P_F) - \frac{\sqrt{N}\mu}{\sigma}\right)$$

The quantity $\frac{\sqrt{N}\mu}{\sigma}$ is called the **signal-to-noise ratio**. As its name suggests, a larger SNR corresponds to improved performance of the Neyman-Pearson decision rule.

Note: In the context of signal processing, the foregoing problem may be viewed as the problem of detecting a constant (DC) signal in additive white Gaussian noise:

$$\mathscr{H}_0 : x_n = w_n, n = 1, \ldots, N$$

$$\mathscr{H}_1 : x_n = A + w_n, n = 1, \ldots, N$$

where A is a known, fixed amplitude, and $w_n \sim \mathcal{N}\left(0, \sigma^2\right)$. Here A corresponds to the mean μ in the example.

The Neyman-Pearson Lemma: General Case

In our initial statement of the Neyman-Pearson Lemma, we assumed that for all η, the set $\{\boldsymbol{x}, \boldsymbol{x} | \ \Lambda(\boldsymbol{x}) = \eta\}$ had probability zero under $\mathscr{H}_0$. This eliminated many important problems from consideration, including tests of discrete data. In this section we remove this restriction.

It is helpful to introduce a more general way of writing decision rules. Let φ be a function of the data $\boldsymbol{x}$ with $\varphi(\boldsymbol{x}) \in [0, 1]$. φ defines the decision rule "declare $\mathscr{H}_1$ with probability $\varphi(\boldsymbol{x})$." In other words, upon observing $\boldsymbol{x}$, we flip a "$\varphi(\boldsymbol{x})$ coin." If it turns up heads, we declare $\mathscr{H}_1$; otherwise we declare $\mathscr{H}_0$. Thus far, we have only considered rules with $\varphi(\boldsymbol{x}) \in \{0, 1\}$
Neyman-Pearson Lemma

Consider the hypothesis testing problem

$$\mathscr{H}_0 : \boldsymbol{x} \sim f_0(\boldsymbol{x})$$

$$\mathscr{H}_1 : \boldsymbol{x} \sim f_1(\boldsymbol{x})$$

where f_0 and f_1 are both pdfs or both pmfs. Let $\alpha \in [0, 1)$ be the size (false-alarm probability) constraint. The decision rule

$$\varphi(\boldsymbol{x}) = \begin{cases} 1 \ \text{ if } \ \Lambda(\boldsymbol{x}) > \eta \\ \rho \ \text{ if } \ \Lambda(\boldsymbol{x}) = \eta \\ 0 \ \text{ if } \ \Lambda(\boldsymbol{x}) < \eta \end{cases}$$

is the most powerful test of size α, where η and ρ are uniquely determined by requiring $P_F = \alpha$. If $\alpha = 0$, we take $\eta = \infty$, $\rho = 0$. This test is unique up to sets of probability zero under $\mathscr{H}_0$ and $\mathscr{H}_1$.
When $\Pr[\Lambda(\boldsymbol{x}) = \eta] > 0$ for certain η, we choose η and ρ as follows: Write

$$P_F = \Pr[\Lambda(\boldsymbol{x}) > \eta] + \rho \Pr[\Lambda(\boldsymbol{x}) = \eta]$$

Choose η such that

$$\Pr[\Lambda(\boldsymbol{x}) > \eta] \leq \alpha \leq \Pr[\Lambda(\boldsymbol{x}) \geq \eta]$$

Then choose ρ such that

$$\rho \Pr[\Lambda(\boldsymbol{x}) = \eta] = \alpha - \Pr[\Lambda(\boldsymbol{x}) < \eta]$$

Example:

Repetition Code

Suppose we have a friend who is trying to transmit a bit (0 or 1) to us over a noisy channel. The channel causes an error in the transmission (that is, the bit is flipped) with probability p, where $0 \leq p < \frac{1}{2}$, and p is known. In order to increase the chance of a successful transmission, our friend sends the same bit N times. Assume the N transmissions are statistically independent. Under these assumptions, the bits you receive are Bernoulli random variables: $x_n \sim \text{Bernoulli}\,(\theta)$. We are faced with the following hypothesis test:

$$\mathscr{H}_0 : \theta = p(0 \text{ sent})$$

$$\mathscr{H}_1 : \theta = 1 - p(1 \text{ sent})$$

We decide to decode the received sequence $\boldsymbol{x} = (x_1 \ldots x_N)^T$ by designing a Neyman-Pearson rule. The likelihood ratio is
Equation:

$$\begin{aligned}\Lambda(\boldsymbol{x}) &= \frac{\prod_{n=1}^{N}(1-p)^{x_n}p^{1-x_n}}{\prod_{n=1}^{N}p^{x_n}(1-p)^{1-x_n}} \\ &= \frac{(1-p)^k p^{N-k}}{p^k(1-p)^{N-k}} \\ &= \left(\frac{1-p}{p}\right)^{2k-N}\end{aligned}$$

where $k = \sum_{n=1}^{N} x_n$ is the number of 1s received.

Note: k is a sufficient statistic for θ.

The LRT is

$$\left(\frac{1-p}{p}\right)^{2k-N} \underset{\mathscr{H}_0}{\overset{\mathscr{H}_1}{\gtrless}} \eta$$

Taking the natural logarithm of both sides and rearranging, we have

$$k \underset{\mathscr{H}_0}{\overset{\mathscr{H}_1}{\gtrless}} \frac{N}{2} + \frac{1}{2}\frac{\ln(\eta)}{\ln\left(\frac{1-p}{p}\right)} = \gamma$$

The false alarm probability is
Equation:

$$\begin{aligned}P_F &= \Pr[k > \gamma] + \rho \Pr[k = \gamma] \\ &= \sum_{k=\gamma+1}^{N}\binom{N}{k}p^k(1-p)^{N-k} + \rho\binom{N}{\gamma}p^{\gamma}(1-p)^{N-\gamma}\end{aligned}$$

γ and ρ are chosen so that $P_F = \alpha$, as described above.

The corresponding detection probability is

Equation:

$$\begin{aligned} P_D &= \Pr[k > \gamma] + \rho \Pr[k = \gamma] \\ &= \sum_{k=\gamma+1}^{N} \binom{N}{k} (1-p)^k p^{N-k} + \rho \binom{N}{\gamma} (1-p)^{\gamma} p^{N-\gamma} \end{aligned}$$

Problems

Exercise:

Problem:

Design a hypothesis testing problem involving continous random variables such that $\Pr[\Lambda(x) = \eta] > 0$ for certain values of η. Write down the false-alarm probability as a function of the threshold. Make as general a statement as possible about when the technical condition is satisfied.

Exercise:

Consider the scalar hypothesis testing problem

$$\mathscr{H}_0 : x \sim f_0(x)$$

$$\mathscr{H}_1 : x \sim f_1(x)$$

where

Problem: $$f_i(x) = \frac{1}{\pi \left(1 + (x-i)^2\right)}, i = \{0, 1\}$$

Write down the likelihood ratio test.

Determine the decision regions as a function of η_1 for all $\eta > 0$. Draw a representative of each. What are the "critical" values of η?

Note: There are five distinct cases.

Compute the size and power (P_F and P_D) in terms of the threshold η_1 and plot the ROC.

Note:

$$\int \frac{1}{1+x^2}\,\mathrm{d}\,x = \arctan(x)$$

Suppose we decide to use a simple threshold test $x \underset{\mathscr{H}_0}{\overset{\mathscr{H}_1}{\gtrless}} \eta$ instead of the Neyman-Pearson rule. Does our performance $\mathscr{H}_0$ suffer much? Plot the ROC for this decision rule on the same graph as for the previous ROC.

Exercise:

Problem:

Suppose we observe N independent realizations of a Poisson random variable k with intensity parameter λ:

$$f(k) = \frac{e^{-\lambda}\lambda^k}{k!}$$

We must decide which of two intensities is in effect:

$$\mathscr{H}_0 : \lambda = \lambda_0$$

$$\mathscr{H}_1 : \lambda = \lambda_1$$

where $\lambda_0 < \lambda_1$.

Write down the likelihood ratio test.

Simplify the LRT to a test statistic involving only a sufficient statistic. Apply a monotonically increasing transformation to simplify further.

Determine the distribution of the sufficient statistic under both hypotheses.

> **Note:**Use the characteristic function to show that a sum of IID Poisson variates is again Poisson distributed.

Derive an expression for the probability of error.

Assuming the two hypotheses are equally likely, and $\lambda_0 = 5$ and $\lambda_1 = 6$, what is the minimum number N of observations needed to attain a false-alarm probability no greater than 0.01?

> **Note:**If you have numerical trouble, try rewriting the log-factorial so as to avoid evaluating the factorial of large integers.

Exercise:

Problem:

In [link], suppose $p = 0.1$. What is the smallest value of N needed to ensure $P_F \leq 0.01$? What is P_D in this case?

White Gaussian Noise

By far the easiest detection problem to solve occurs when the noise vector consists of statistically independent, identically distributed, Gaussian random variables. In this book, a **white sequence** consists of statistically independent random variables. The white sequence's mean is usually taken to be zero [footnote] and each component's variance is σ^2. The equal-variance assumption implies the noise characteristics are unchanging throughout the entire set of observations. The probability density of the zero-mean noise vector evaluated at $\boldsymbol{r} - s_i$ equals that of Gaussian random vector having independent components ($K = \sigma^2 I$) with mean s_i .

$$\mathrm{p}_{\boldsymbol{n}}\left(\boldsymbol{r} - \boldsymbol{s}_i\right) = \left(\frac{1}{2\pi\sigma^2}\right)^{\frac{L}{2}} e^{-\left(\frac{1}{2\sigma^2}(\boldsymbol{r}-s_i)^T(\boldsymbol{r}-s_i)\right)}$$

The resulting detection problem is similar to the Gaussian example examined so frequently in the hypothesis testing sections, with the distinction here being a non-zero mean under both models. The logarithm of the likelihood ratio becomes

$$\left(\boldsymbol{r} - s_0\right)^T\left(\boldsymbol{r} - s_0\right) - \left(\boldsymbol{r} - s_1\right)^T\left(\boldsymbol{r} - s_1\right) \underset{\mathscr{M}_0}{\overset{\mathscr{M}_1}{\gtrless}} 2\sigma^2 \ln(\eta)$$

and the usual simplifications yield in

$$\boldsymbol{r}^T s_1 - \frac{{s_1}^T s_1}{2} - \left(\boldsymbol{r}^T s_0 - \frac{{s_0}^T s_0}{2}\right) \underset{\mathscr{M}_0}{\overset{\mathscr{M}_1}{\gtrless}} \sigma^2 \ln(\eta)$$

The quantities in parentheses express the signal processing operations for each model. If more than two signals were assumed possible, quantities such as these would need to be computed for each signal and the largest selected. This decision rule is optimum for the additive, white Gaussian noise problem.
The zero-mean assumption is realistic for the detection problem. If the mean were non-zero, simply subtracting it from the observed sequence results in a zero-mean noise component.

Each term in the computations for the optimum detector has a signal processing interpretation. When expanded, the term ${s_i}^T s_i$ equals $\sum_{l=0}^{L-1} {s_i}^2(l)$, which is the **signal energy** E_i. The remaining term - $\boldsymbol{r}^T s_i$ - is the only one involving the observations and hence constitutes the sufficient statistic $\Upsilon_i(\boldsymbol{r})$ for the additive white Gaussian noise detection problem.

$$\Upsilon_i(\boldsymbol{r}) = \boldsymbol{r}^T s_i$$

An abstract, but physically relevant, interpretation of this important quantity comes from the theory of linear vector spaces. There, the quantity $\boldsymbol{r}^T s_i$ would be termed the **dot product** between $\boldsymbol{r}$ and s_i or the **projection** of $\boldsymbol{r}$ onto s_i . By employing the Schwarz inequality, the largest value of this quantity occurs when these vectors are proportional to each other. Thus, a dot product computation measures how much alike two vectors are: they are completely alike when they are parallel (proportional) and completely dissimilar when orthogonal (the dot product is zero). More precisely, the dot product removes those components from the observations which are orthogonal to the signal. The dot product thereby generalizes the familiar notion of filtering a signal contaminated by broadband noise. In filtering, the signal-to-noise ratio of a bandlimited signal can be drastically improved by lowpass filtering; the output would consist only of the signal and "in-band" noise. The dot product serves a similar role, ideally removing those "out-of-band" components (the orthogonal ones) and retaining the "in-band" ones (those parallel to the signal).

Expanding the dot product, $\boldsymbol{r}^T s_i = \sum_{l=0}^{L-1} r(l)s_i(l)$ another signal processing interpretation emerges. The dot product now describes a finite impulse response (FIR) filtering operation evaluated at a specific index. To demonstrate this interpretation, let $h(l)$ be the unit-sample response of a linear, shift-invariant filter where $h(l) = 0$ for $l < 0$ and $l \geq L$. Letting $r(l)$ be the filter's input sequence, the convolution sum expresses the output.

$$r(k)*h(k) = \sum_{l=k-(L-1)}^{k} r(l)h(k-l)$$

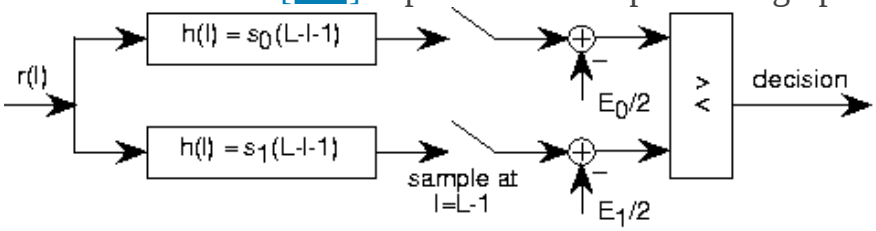
Letting $k = L - 1$, the index at which the unit-sample response's last value overlaps the input's value at the origin, we have

$$r(k)*h(k)|_{k=L-1} = \sum_{l=0}^{L-1} r(l)h(L-1-l)$$

If we set the unit-sample response equal to the index-reversed, then delayed signal $h(l) = s_i(L-1-l)$, we have

$$r(k)*s_i(L-1-k)|_{k=L-1} = \sum_{l=0}^{L-1} r(l)s_i(l)$$

which equals the observation-dependent component of the optimal detector's sufficient statistic. [link] depicts these computations graphically.



The detector for signals contained in additive, white Gaussian noise consists of a matched filter, whose output is sampled at the duration of the signal and half of the signal energy is subtracted from it. The optimum detector incorporates a matched filter for each signal compares their outputs to determine the largest.

The sufficient statistic for the i^{th} signal is thus expressed in signal processing notation as $r(k)*s_i(L-1-k)|_{k=L-1} - \frac{E_i}{2}$. The filtering term is called a **matched filter** because the observations are passed through a filter whose unit-sample response "matches" that of the signal being sought. We sample the matched filter's output at the precise moment when all of the observations fall within the filter's memory and then adjust this value by half the signal energy. The adjusted values for the two assumed signals are subtracted and compared to a threshold.

To compute the performance probabilities, the expressions should be simplified in the ways discussed in the hypothesis testing sections. As the energy terms are known a priori they can be incorporated into the threshold with the result

$$\sum_{l=0}^{L-1} r(l)\left(s_1(l)-s_0(l)\right) \underset{\mathscr{M}_0}{\overset{\mathscr{M}_1}{\gtrless}} \sigma^2 \ln(\eta) + \frac{E_1 - E_0}{2}$$

The left term constitutes the sufficient statistic for the binary detection problem. Because the additive noise is presumed Gaussian, the sufficient statistic is a Gaussian random variable no matter which model is assumed. Under $\mathscr{M}_i$, the specifics of this probability distribution are

$$\sum_{l=0}^{L-1} r(l)\left(s_1(l)-s_0(l)\right) \sim \mathcal{N}\left(\sum s_i(l)\left(s_1(l)-s_0(l)\right), \sigma^2 \sum \left(s_1(l)-s_0(l)\right)^2\right)$$

The false-alarm probability is given by

$$P_F = Q\left(\frac{\sigma^2 \ln(\eta) + \frac{E_1-E_0}{2} - \sum s_0(l)\left(s_1(l)-s_0(l)\right)}{\sigma\left(\sum \left(s_1(l)-s_0(l)\right)^2\right)^{\frac{1}{2}}}\right)$$

The signal-related terms in the numerator of this expression can be manipulated with the false-alarm probability (and the detection probability) for the optimal white Gaussian noise detector succinctly expressed by

$$P_F = Q\left(\frac{\ln(\eta) + \frac{1}{2\sigma^2}\sum \left(s_1(l)-s_0(l)\right)^2}{\frac{1}{\sigma}\left(\sum \left(s_1(l)-s_0(l)\right)^2\right)^{\frac{1}{2}}}\right)$$

$$P_F = Q\left(\frac{\ln(\eta) - \frac{1}{2\sigma^2}\sum \left(s_1(l)-s_0(l)\right)^2}{\frac{1}{\sigma}\left(\sum \left(s_1(l)-s_0(l)\right)^2\right)^{\frac{1}{2}}}\right)$$

Note that the **only** signal-related quantity affecting this performance probability (and all of the others) is the **ratio of energy in the difference signal to the noise variance.** The larger this ratio, the better (smaller) the performance probabilities become. Note that the details of the signal waveforms do not greatly affect the energy of the difference signal. For example, consider the case where the two signal energies are equal ($E_0 = E_1 = E$); the energy of the difference signal is given by $2E - 2\sum s_0(l)s_1(l)$. The largest value of this energy occurs when the signals are negatives of each other, with the difference-signal energy equaling $4E$. Thus, equal-energy but opposite-signed signals such as sine waves, square-waves, Bessel functions, etc. **all** yield exactly the same performance levels. The essential signal properties that do yield good performance values are elucidated by an alternate interpretation. The term $\sum \left(s_1(l)-s_0(l)\right)^2$ equals $\left(\| s_1 - s_0 \|\right)^2$, the L^2 norm of the difference signal. Geometrically, the difference-signal energy is the same quantity as the square of the Euclidean distance between the two signals. In

these terms, a larger distance between the two signals will mean better performance.

Example:
Detection, Gaussian example
A common detection problem in array processing is to determine whether a signal is present ($\mathscr{M}_1$) or not ($\mathscr{M}_0$) in the array output. In this case, $s_0(l) = 0$ The optimal detector relies on filtering the array output with a matched filter having an impulse response based on the assumed signal. Letting the signal under $\mathscr{M}_1$ be denoted simply by $s(l)$, the optimal detector consists of

$$r(l)*s(L-1-l)|_{l=L-1} - \frac{E}{2} \underset{\mathscr{M}_0}{\overset{\mathscr{M}_1}{\gtrless}} \sigma^2 \ln(\eta)$$

or

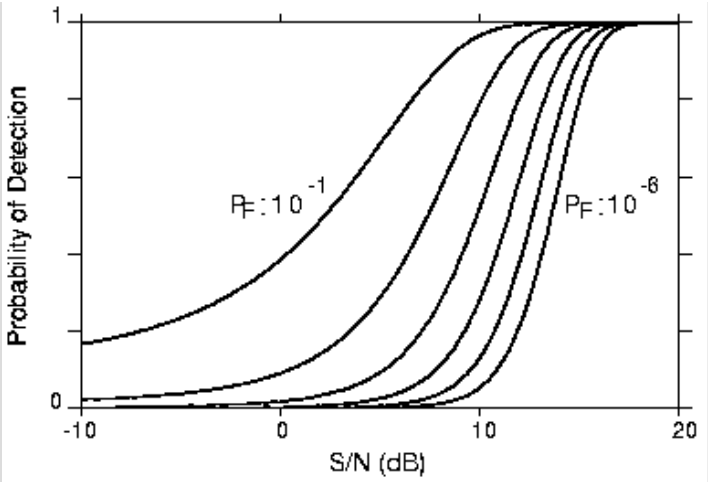
$$r(l)*s(L-1-l)|_{l=L-1} \underset{\mathscr{M}_0}{\overset{\mathscr{M}_1}{\gtrless}} \gamma$$

The false-alarm and detection probabilities are given by

$$P_F = Q\left(\frac{\gamma}{\frac{E^{\frac{1}{2}}}{\sigma}}\right)$$

$$P_D = Q\left(Q^{-1}(P_F) - \sqrt{\frac{E}{\sigma}}\right)$$

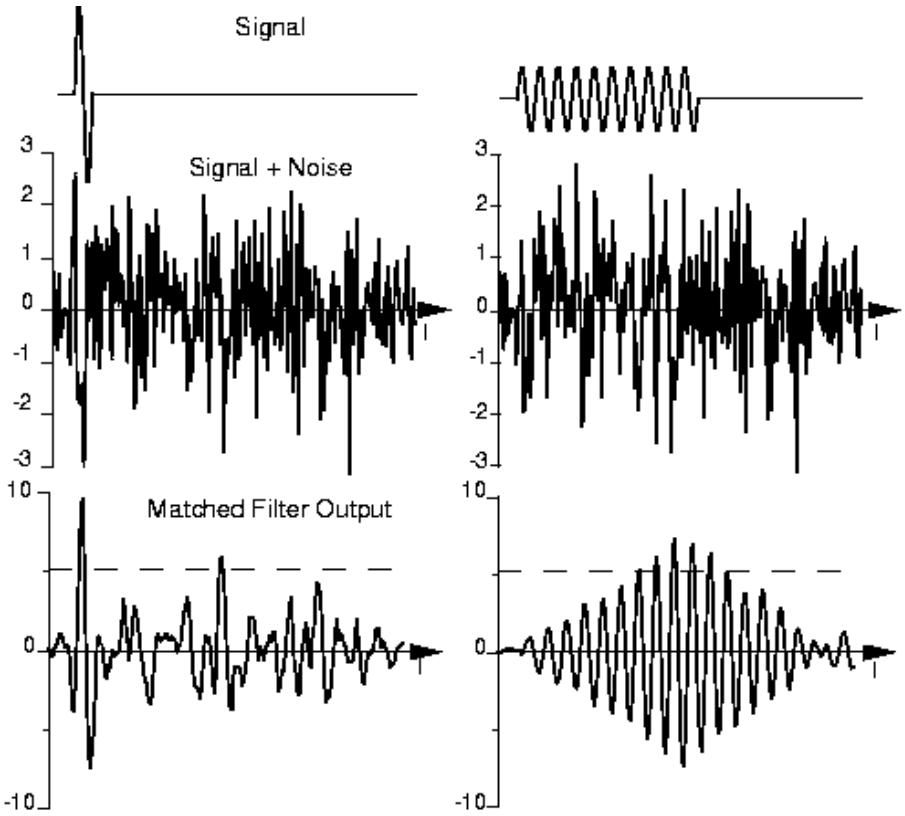
[link] displays the probability of detection as a function of the signal-to-noise ratio $\frac{E}{\sigma^2}$ for several values of false-alarm probability. Given an estimate of the expected signal-to-noise ratio, these curves can be used to assess the trade-off between the false-alarm and detection probabilities.



The probability of detection is plotted versus signal-to-noise ratio for various values of the false-alarm probability P_F. False-alarm probabilities range from 10^{-1} down to 10^{-6} by decades. The matched filter receiver was used since the noise is white and Gaussian. Note how the range of signal-to-noise ratios over which the detection probability changes shrinks as the false-alarm probability decreases. This effect is a consequence of the non-linear nature of the function $Q(\cdot)$.

The important parameter determining detector performance derived in this example is the **signal-to-noise ratio** $\frac{E}{\sigma^2}$: the larger it is, the smaller the false-alarm probability is (generally speaking). Signal-to-noise ratios can be measured in many different ways. For example, one measure might be the ratio of the rms signal amplitude to the rms noise amplitude. Note that the important one for the detection problem is much different. The signal portion is the **sum** of the squared signal values over the **entire** set of observed values - the signal energy; the noise portion is the variance of **each** noise component - the noise power. Thus, energy can be increased in two ways that increase the signal-to-noise ratio: the signal can be made larger **or** the observations can be extended to encompass a larger number of values.

To illustrate this point, two signals having the same energy are shown in [link]. When these signals are shown in the presence of additive noise, the signal is visible on the left because its amplitude is larger; the one on the right is much more difficult to discern. The instantaneous signal-to-noise ratio-the ratio of signal amplitude to average noise amplitude - is the important visual cue. However, the kind of signal-to-noise ratio that determines detection performance belies the eye. The matched filter outputs have similar maximal values, indicating that total signal energy rather than amplitude determines the performance of a matched filter detector.



Two signals having the same energy are shown at the top of the figure. The one on the left equals one cycle of a sinusoid having ten samples/period ($\sin(\omega_0 l)$ with $\omega_0 = 2\pi 0.1$). On the right, ten cycles of similar signal is shown, with an amplitude a factor of $\sqrt{10}$ smaller. The middle portion of the figure shows these signals with the same noise signal added; the duration of this signal is 200 samples. The lower portion depicts the outputs of matched filters for each signal. The detection threshold was set by specifying a false-alarm probability of 10^{-2}.

Validity of the White Noise Assumption

The optimal detection paradigm for the additive, white Gaussian noise problem has a relatively simple solution: construct FIR filters whose unit-sample responses are related to the presumed signals and compare the filtered outputs with a threshold. We may well wonder which assumptions made in this problem are most questionable in "real-world" applications. noise is additive in most cases. In many situation, the additive noise present in observed data is Gaussian. Because of the Central Limit Theorem, if numerous noise sources impinge on a measuring device, their superposition will be Gaussian to a great extent. As we know from the discussion on the Central Limit Theorem, glibly appealing to the Central Limit Theorem is not without hazards; the non-Gaussian detection problem will be discussed in some detail later. Interestingly, the weakest assumption is the "whiteness" of the noise. Note that the observation sequence is obtained as a result of **sampling** the sensor outputs. Assuming white noise samples does **not** mean that the continuous-time noise was white. White noise in continuous time has infinite variance and cannot be sampled; discrete-time white noise has a finite variance with a constant power spectrum. The Sampling Theorem suggests that a signal is represented accurately by its samples only if we choose a sampling frequency commensurate with the signal's bandwidth. One should note that fidelity of representation does **not** mean that the sample values are independent. In most cases, satisfying the Sampling Theorem means that the samples are correlated. As shown in Sampling and Random Sequences, the correlation function of sampled noise equals samples of the original correlation function. For the sampled noise to be white, $E[n(l_1T)n(l_2T)] = 0$ for $l_1 \neq l_2$: the samples of the correlation function at locations other than the origin must all be zero. While some correlation functions have this property, **many examples satisfy the sampling theorem but do not yield uncorrelated samples**. In many practical situations, **undersampling** the noise will reduce inter-sample correlation. Thus, we obtain uncorrelated samples either by deliberately undersampling, which wastes signal energy, or by imposing anti-aliasing filters that have a bandwidth larger than the signal and sampling at the signal's Nyquist rate. Since the noise power spectrum usually extends to higher frequencies than the signal, this intentional undersampling can result in larger noise variance. in either case, by trying to make the problem at hand match the solution, we are actually reducing performance! We need a **direct** approach to attacking the correlated noise issue that arises in virtually **all** sampled-data detection problems rather than trying to work around it.

Elementary Hypothesis Testing

In statistics, hypothesis testing is some times known as decision theory or simply testing. The key result around which all decision theory revolves is the likelihood ratio test.

The Likelihood Ratio Test

In a binary hypothesis testing problem, four possible outcomes can result. Model $\mathscr{M}_0$ did in fact represent the best model for the data and the decision rule said it was (a correct decision) or said it wasn't (an erroneous decision). The other two outcomes arise when model $\mathscr{M}_1$ was in fact true with either a correct or incorrect decision made. The decision process operates by segmenting the range of observation values into two disjoint **decision regions** $\Re_0$ and $\Re_1$. All values of $\boldsymbol{r}$ fall into either $\Re_0$ or $\Re_1$. If a given $\boldsymbol{r}$ lies in $\Re_0$, for example, we will announce our decision ""model $\mathscr{M}_0$ was true""; if in $\Re_1$, model $\mathscr{M}_1$ would be proclaimed. To derive a rational method of deciding which model best describes the observations, we need a criterion to assess the quality of the decision process. Optimizing this criterion will specify the decision regions.

The **Bayes' decision criterion** seeks to minimize a cost function associated with making a decision. Let C_{ij} be the cost of mistaking model j for model i ($i \neq j$) and C_{ii} the presumably smaller cost of correctly choosing model i: $C_{ij} > C_{ii}$, $i \neq j$. Let π_i be the a priori probability of model i. The so-

called **Bayes' cost** C is the average cost of making a decision.
Equation:

$$\begin{aligned} C &= \textstyle\sum_{i,j\{i,j\}} C_{ij}\pi_j \Pr[\text{say } \mathscr{M}_i \text{ when } H_j \text{ true}] \\ &= \textstyle\sum_{i,j\{i,j\}} C_{ij}\pi_j \Pr[\text{say } \mathscr{M}_i \mid H_j \text{ true}] \end{aligned}$$

The Bayes' cost can be expressed as
Equation:

$$\begin{aligned} C &= \textstyle\sum_{i,j\{i,j\}} C_{ij}\pi_j \Pr[\boldsymbol{r} \in \Re_i \mid \mathscr{M}_0 \text{ true}] \\ &= \textstyle\sum_{i,j\{i,j\}} C_{ij}\pi_j \int \mathrm{p}_{\boldsymbol{r}\mid H_j}(\boldsymbol{r})\,\mathrm{d}\,\boldsymbol{r} \\ &= \int C_{00}\pi_0\,\mathrm{p}_{\boldsymbol{r}\mid\mathscr{M}_0}(\boldsymbol{r}) + C_{01}\pi_1\,\mathrm{p}_{\boldsymbol{r}\mid\mathscr{M}_1}(\boldsymbol{r})\,\mathrm{d}\,\boldsymbol{r} + \int C_{10}\pi_0\,\mathrm{p}_{\boldsymbol{r}\mid\mathscr{M}_0}(\boldsymbol{r}) + C_{11}\pi_1\,\mathrm{p}_{\boldsymbol{r}\mid\mathscr{M}_1}(\boldsymbol{r})\,\mathrm{d}\,\boldsymbol{r} \end{aligned}$$

$\mathrm{p}_{\boldsymbol{r}\mid\mathscr{M}_i}(\boldsymbol{r})$ is the conditional probability density function of the observed data $\boldsymbol{r}$ given that model $\mathscr{M}_i$ was true. To minimize this expression with respect to the decision regions $\Re_0$ and $\Re_1$, ponder which integral would yield the smallest value if its integration domain included a specific observation vector. This selection process defines the decision regions; for example, we choose $\mathscr{M}_0$ for those values of $\boldsymbol{r}$ which yield a smaller value for the first integral.

$$\pi_0 C_{00}\,\mathrm{p}_{\boldsymbol{r}\mid\mathscr{M}_0}(\boldsymbol{r}) + \pi_1 C_{01}\,\mathrm{p}_{\boldsymbol{r}\mid\mathscr{M}_1}(\boldsymbol{r}) < \pi_0 C_{10}\,\mathrm{p}_{\boldsymbol{r}\mid\mathscr{M}_0}(\boldsymbol{r}) + \pi_1 C_{11}\,\mathrm{p}_{\boldsymbol{r}\mid\mathscr{M}_1}(\boldsymbol{r})$$

We choose $\mathscr{M}_1$ when the inequality is reversed. This expression is easily manipulated to obtain the decision rule known as the **likelihood ratio test**.
Equation:

$$\frac{\mathrm{p}_{\boldsymbol{r}|\mathscr{M}_1}(\boldsymbol{r})}{\mathrm{p}_{\boldsymbol{r}|\mathscr{M}_0}(\boldsymbol{r})} \underset{\mathscr{M}_0}{\overset{\mathscr{M}_1}{\gtrless}} \frac{\pi_0(C_{10}-C_{00})}{\pi_1(C_{01}-C_{11})}$$

The comparison relation means selecting model $\mathscr{M}_1$ if the left-hand ratio exceeds the value on the right; otherwise, $\mathscr{M}_0$ is selected. Thus, the **likelihood ratio** $\frac{\mathrm{p}_{\boldsymbol{r}|\mathscr{M}_1}(\boldsymbol{r})}{\mathrm{p}_{\boldsymbol{r}|\mathscr{M}_0}(\boldsymbol{r})}$ symbolically represented by $\Lambda(\boldsymbol{r})$, is computed from the observed value of $\boldsymbol{r}$ and then compared with a **threshold** η equaling $\frac{\pi_0(C_{10}-C_{00})}{\pi_1(C_{01}-C_{11})}$. Thus, when two models are hypothesized, the likelihood ratio test can be succinctly expressed as the comparison of the likelihood ratio with a threshold.
Equation:

$$\Lambda(\boldsymbol{r}) \underset{\mathscr{M}_0}{\overset{\mathscr{M}_1}{\gtrless}} \eta$$

The data processing operations are captured entirely by the likelihood ratio $\frac{\mathrm{p}_{\boldsymbol{r}|\mathscr{M}_1}(\boldsymbol{r})}{\mathrm{p}_{\boldsymbol{r}|\mathscr{M}_0}(\boldsymbol{r})}$. Furthermore, note that only the value of the likelihood ratio **relative** to the threshold matters; to simplify the computation of the likelihood ratio, we can perform **any** positively monotonic operations simultaneously on the likelihood ratio and the threshold without affecting the comparison. We can multiply the ratio by a positive constant, add any constant, or apply a monotonically increasing function which simplifies the expressions. We single one such function, the logarithm, because it simplifies likelihood ratios that commonly occur in signal processing applications. Known as the log-likelihood, we explicitly express the likelihood ratio test with it as
Equation:

$$\ln(\Lambda(\boldsymbol{r})) \underset{\mathscr{M}_0}{\overset{\mathscr{M}_1}{\gtrless}} \ln(\eta)$$

Useful simplifying transformations are problem-dependent; by laying bare that aspect of the observations essential to the model testing problem, we reveal the **sufficient statistic** $\Upsilon(\boldsymbol{r})$: the scalar quantity which best summarizes the data (Lehmann, pp. 18-22). The likelihood ratio test is best expressed in terms of the sufficient statistic.
Equation:

$$\Upsilon(\boldsymbol{r}) \underset{\mathscr{M}_0}{\overset{\mathscr{M}_1}{\gtrless}} \gamma$$

We will denote the threshold value by γ when the sufficient statistic is used or by η when the likelihood ratio appears prior to its reduction to a sufficient statistic.

As we shall see, if we use a different criterion other than the Bayes' criterion, the decision rule often involves the likelihood ratio. The likelihood ratio is comprised of the quantities $\mathrm{p}_{\boldsymbol{r}|\mathscr{M}_i}(\boldsymbol{r})$, termed the **likelihood function**, which is also important in estimation theory. It is this conditional density that portrays the probabilistic model describing data generation. The likelihood function completely

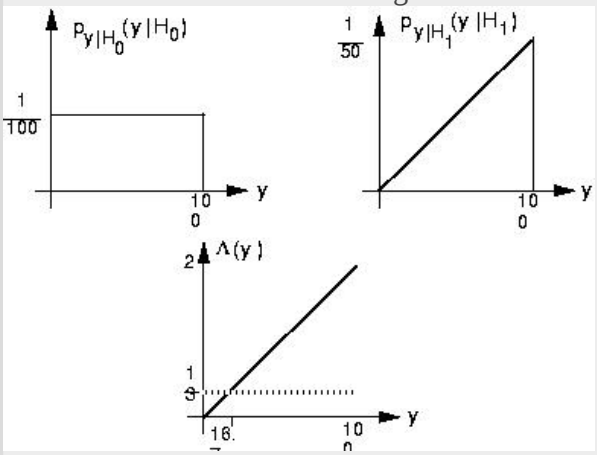
characterizes the kind of "world" assumed by each model; for each model, we must specify the likelihood function so that we can solve the hypothesis testing problem.

A complication, which arises in some cases, is that the sufficient statistic may not be monotonic. If monotonic, the decision regions $\mathfrak{R}_0$ and $\mathfrak{R}_1$ are simply connected (all portions of a region can be reached without crossing into the other region). If not, the regions are not simply connected and decision region islands are created (see this problem). Such regions usually complicate calculations of decision performance. Monotonic or not, the decision rule proceeds as described: the sufficient statistic is computed for each observation vector and compared to a threshold.

Example:
An instructor in a course in detection theory wants to determine if a particular student studied for his last test. The observed quantity is the student's grade, which we denote by $\boldsymbol{r}$. Failure may not indicate studiousness: conscientious students may fail the test. Define the models as

- $\mathscr{M}_0$: did not study
- $\mathscr{M}_1$: did study

The conditional densities of the grade are shown in [link].



Conditional densities for the grade distributions assuming that a student did not study ($\mathscr{M}_0$) or did ($\mathscr{M}_1$) are shown in the top row. The lower portion depicts the likelihood ratio formed from these densities.

Based on knowledge of student behavior, the instructor assigns a priori probabilities of $\pi_0 = 1/4$ and $\pi_1 = 3/4$. The costs C_{ij} are chosen to reflect the instructor's sensitivity to student feelings: $C_{01} = 1 = C_{10}$ (an erroneous decision either way is given the same cost) and $C_{00} = 0 = C_{11}$. The likelihood ratio is plotted in [link] and the threshold value η, which is computed from the a priori probabilities and the costs to be $1/3$, is indicated. The calculations of this comparison can be simplified in an obvious way.

$$\frac{r}{50} \underset{\mathscr{M}_0}{\overset{\mathscr{M}_1}{\gtrless}} 1/3$$

or

$$r \underset{\mathscr{M}_0}{\overset{\mathscr{M}_1}{\gtrless}} 50/3 = 16.7$$

The multiplication by the factor of 50 is a simple illustration of the reduction of the likelihood ratio to a sufficient statistic. Based on the assigned costs and a priori probabilities, the optimum decision rule says the instructor must assume that the student did not study if the student's grade is less than 16.7; if greater, the student is assumed to have studied despite receiving an abysmally low grade such as 20. Note that as the densities given by each model overlap entirely: the possibility of making the wrong interpretation **always** haunts the instructor. However, no other procedure will be better!

The Bayes Risk Criterion in Hypothesis Testing

The design of a hypothesis test/detector often involves constructing the solution to an optimization problem. The optimality criteria used fall into two classes: Bayesian and frequent.

In the Bayesian setup, it is assumed that the a priori probability of each hypothesis occuring (π_i) is known. A cost C_{ij} is assigned to each possible outcome:

$$C_{\mathrm{ij}} = \Pr[\mathrm{say} H_i \mathrm{when} H_j \mathrm{true}]$$

The optimal test/detector is the one that minimizes the Bayes risk, which is defined to be the expected cost of an experiment:

$$C = \sum_{i,j} C_{\mathrm{ij}} \pi_i \Pr[\mathrm{say} H_i \mathrm{when} H_j \mathrm{true}]$$

In the event that we have a binary problem, and both hypotheses are simple, the decision rule that minimizes the Bayes risk can be constructed explicitly. Let us assume that the data is continuous (i.e., it has a density) under each hypothesis:

$$H_0 : \boldsymbol{x} \sim f_0(\boldsymbol{x})$$

$$H_1 : \boldsymbol{x} \sim f_1(\boldsymbol{x})$$

Let R_0 and R_1 denote the decision regions corresponding to the optimal test. Clearly, the optimal test is specified once we specify R_0 and $R_1 = {R_0}'$.

The Bayes risk may be written
Equation:

$$\begin{aligned} \bar{C} &= \textstyle\sum_{\{i,j\}=0}^{1} C_{ij} \pi_i \int f_j(\boldsymbol{x}) \, \mathrm{d}\, \boldsymbol{x} \\ &= \int C_{00} \pi_0 f_0(\boldsymbol{x}) + C_{01} \pi_1 f_1(\boldsymbol{x}) \, \mathrm{d}\, \boldsymbol{x} + \int C_{10} \pi_0 f_0(\boldsymbol{x}) + C_{11} \pi_1 f_1(\boldsymbol{x}) \, \mathrm{d}\, \boldsymbol{x} \end{aligned}$$

Recall that R_0 and R_1 **partition** the input space: they are disjoint and their union is the full input space. Thus, every possible input $\boldsymbol{x}$ belongs to precisely one of these regions. In order to minimize the Bayes risk, a measurement $\boldsymbol{x}$ should belong to the decision region R_i for which the corresponding integrand in the

preceding equation is smaller. Therefore, the Bayes risk is minimized by assigning $\boldsymbol{x}$ to R_0 whenever

$$\pi_0 C_{00} f_0(\boldsymbol{x}) + \pi_1 C_{01} f_1(\boldsymbol{x}) < \pi_0 C_{10} f_0(\boldsymbol{x}) + \pi_1 C_{11} f_1(\boldsymbol{x})$$

and assigning $\boldsymbol{x}$ to R_1 whenever this inequality is reversed. The resulting rule may be expressed concisely as

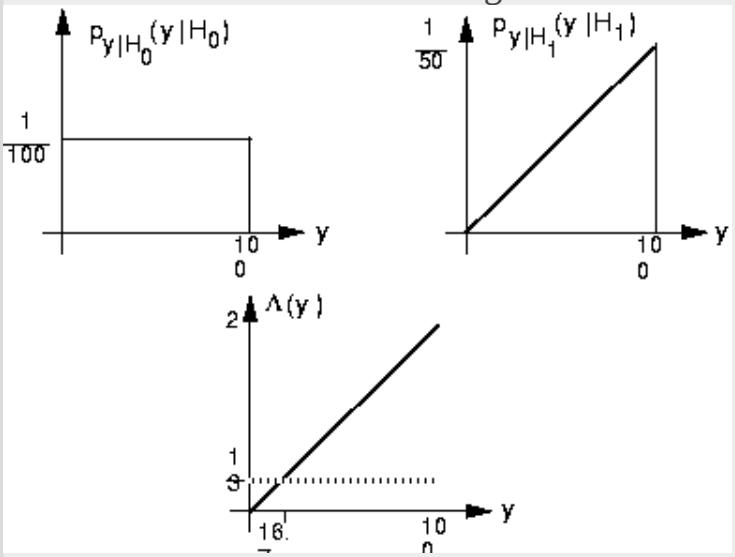
$$\Lambda(\boldsymbol{x}) \equiv \frac{f_1(\boldsymbol{x})}{f_0(\boldsymbol{x})} \underset{H_0}{\overset{H_1}{\gtrless}} \frac{\pi_0 \left(C_{10} - C_{00}\right)}{\pi_1 \left(C_{01} - C_{11}\right)} \equiv \eta$$

Here, $\Lambda(\boldsymbol{x})$ is called the **likelihood ratio**, η is called the threshold, and the overall decision rule is called the Likelihood Ratio Test (LRT). The expression on the right is called a **threshold**.

Example:
An instructor in a course in detection theory wants to determine if a particular student studied for his last test. The observed quantity is the student's grade, which we denote by $\boldsymbol{r}$. Failure may not indicate studiousness: conscientious students may fail the test. Define the models as

- $_0$: did not study
- $_1$: did study

The conditional densities of the grade are shown in [link].



Conditional densities for the grade distributions assuming that a student did not study ($_0$) or did ($_1$) are shown in the top row. The lower portion depicts the likelihood ratio formed from these densities.

Based on knowledge of student behavior, the instructor assigns a priori probabilities of $\pi_0 = \frac{1}{4}$ and $\pi_1 = \frac{3}{4}$. The costs C_{ij} are chosen to reflect the instructor's sensitivity to student feelings: $C_{01} = 1 = C_{10}$ (an erroneous decision either way is given the same cost) and $C_{00} = 0 = C_{11}$. The likelihood ratio is plotted in [link] and the threshold value η, which is computed from the a priori probabilities and the costs to be $\frac{1}{3}$, is indicated. The calculations of this comparison can be simplified in an obvious way.

$$\frac{r}{50} \underset{0}{\overset{1}{\gtrless}} \frac{1}{3}$$

or

$$r \underset{0}{\overset{1}{\gtrless}} \frac{50}{3} = 16.7$$

The multiplication by the factor of 50 is a simple illustration of the reduction of the likelihood ratio to a sufficient statistic. Based on the assigned costs and a priori probabilities, the optimum decision rule says the instructor must assume that the student did not study if the student's grade is less than 16.7; if greater, the student is assumed to have studied despite receiving an abysmally low grade such as 20. Note that as the densities given by each model overlap entirely: the possibility of making the wrong interpretation **always** haunts the instructor. However, no other procedure will be better!

A special case of the minimum Bayes risk rule, the minimum probability of error rule, is used extensively in practice, and is discussed at length in another module.

Problems

Exercise:

Problem:

Denote $\alpha = \Pr[\text{declare} H_1 \text{when} H_0 \text{true}]$ and

$\beta = \Pr[\text{declare} H_1 \text{when} H_1 \text{true}]$. Express the Bayes risk $\bar{C}$ in terms of α and β, C_{ij}, and π_i. Argue that the optimal decision rule is not altered by setting $C_{00} = C_{11} = 0$.

Exercise:

Problem:

Suppose we observe $\boldsymbol{x}$ such that $\Lambda(\boldsymbol{x}) = \eta$. Argue that it doesn't matter whether we assign $\boldsymbol{x}$ to R_0 or R_1.

Sufficient Statistics

Introduction

Sufficient statistics arise in nearly every aspect of statistical inference. It is important to understand them before progressing to areas such as hypothesis testing and parameter estimation.

Suppose we observe an N-dimensional random vector $\boldsymbol{X}$, characterized by the density or mass function $\mathrm{f}_{\boldsymbol{\theta}}\left(\boldsymbol{x}\right)$, where $\boldsymbol{\theta}$ is a p-dimensional vector of parameters to be estimated. The functional form of $\mathrm{f}\left(x\right)$ is assumed known. The parameter $\boldsymbol{\theta}$ completely determines the distribution of $\boldsymbol{X}$. Conversely, a measurement $\boldsymbol{x}$ of $\boldsymbol{X}$ provides information about $\boldsymbol{\theta}$ through the probability law $\mathrm{f}_{\boldsymbol{\theta}}\left(\boldsymbol{x}\right)$.

Example:
Suppose $\boldsymbol{X} = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}$, where $X_i \sim \mathcal{N}(\theta, 1)$ are IID. Here θ is a scalar parameter specifying the mean. The distribution of $\boldsymbol{X}$ is determined by θ through the density

$$\mathrm{f}_{\theta}\left(\boldsymbol{x}\right) = \frac{1}{\sqrt{2\pi}} e^{-\frac{\left(x_1-\theta\right)^2}{2}} \frac{1}{\sqrt{2\pi}} e^{-\frac{\left(x_2-\theta\right)^2}{2}}$$

On the other hand, if we observe $\boldsymbol{x} = \begin{pmatrix} 100 \\ 102 \end{pmatrix}$, then we may safely assume $\theta = 0$ is highly unlikely.

The N-dimensional observation $\boldsymbol{X}$ carries information about the p-dimensional parameter vector $\boldsymbol{\theta}$. If $p < N$, one may ask the following question: Can we compress $\boldsymbol{x}$ into a low-dimensional statistic without any loss of information? Does there exist some function $\boldsymbol{t} = T(\boldsymbol{x})$, where the

dimension of $\boldsymbol{t}$ is $M < N$, such that $\boldsymbol{t}$ carries all the useful information about $\boldsymbol{\theta}$?

If so, for the purpose of studying $\boldsymbol{\theta}$ we could discard the raw measurements $\boldsymbol{x}$ and retain only the low-dimensional statistic $\boldsymbol{t}$. We call $\boldsymbol{t}$ a **sufficient statistic**. The following definition captures this notion precisely:

> Let $X_1, \ldots, X_M$ be a random sample, governed by the density or probability mass function $\mathrm{f}\ (\boldsymbol{x} \,|\, \boldsymbol{\theta})$. The statistic $T(\boldsymbol{x})$ is **sufficient** for $\boldsymbol{\theta}$ if the conditional distribution of $\boldsymbol{x}$, given $T(\boldsymbol{x}) = \boldsymbol{t}$, is independent of $\boldsymbol{\theta}$. Equivalently, the functional form of $\mathrm{f}_{\boldsymbol{\theta} \,|\, \boldsymbol{t}}\ (\boldsymbol{x})$ does not involve $\boldsymbol{\theta}$.

How should we interpret this definition? Here are some possibilities:

1. Let $\mathrm{f}_{\boldsymbol{\theta}}\ (\boldsymbol{x}, \boldsymbol{t})$ denote the joint density or probability mass function on $(\boldsymbol{X}, \boldsymbol{T}(\boldsymbol{X}))$. If $T(\boldsymbol{X})$ is a sufficient statistic for $\boldsymbol{\theta}$, then
Equation:

$$\begin{aligned} \mathrm{f}_{\boldsymbol{\theta}}\ (\boldsymbol{x}) &= \mathrm{f}_{\boldsymbol{\theta}}\ (\boldsymbol{x}, T(\boldsymbol{x})) \\ &= \mathrm{f}_{\boldsymbol{\theta} \,|\, \boldsymbol{t}}\ (\boldsymbol{x})\ \mathrm{f}_{\boldsymbol{\theta}}\ (\boldsymbol{t}) \\ &= \mathrm{f}\ (\boldsymbol{x} \,|\, \boldsymbol{t})\ \mathrm{f}_{\boldsymbol{\theta}}\ (\boldsymbol{t}) \end{aligned}$$

Therefore, the parametrization of the probability law for the measurement $\boldsymbol{x}$ is manifested in the parametrization of the probability law for the statistic $T(\boldsymbol{x})$.

2. Given $\boldsymbol{t} = T(\boldsymbol{x})$, full knowledge of the measurement $\boldsymbol{x}$ brings no additional information about $\boldsymbol{\theta}$. Thus, we may discard $\boldsymbol{x}$ and retain on the compressed statistic $\boldsymbol{t}$.

3. Any inference strategy based on $\mathrm{f}_{\boldsymbol{\theta}}\left(\boldsymbol{x}\right)$ may be replaced by a strategy based on $\mathrm{f}_{\boldsymbol{\theta}}\left(\boldsymbol{t}\right)$.

Example:

Binary Information Source

(Scharf, pp.78) Suppose a binary information source emits a sequence of binary (0 or 1) valued, independent variables $x_1, \ldots, x_N$. Each binary symbol may be viewed as a realization of a Bernoulli trial: $x_n \sim \text{Bernoulli}\left(\theta\right)$, iid. The parameter $\theta \in [0, 1]$ is to be estimated.

The probability mass function for the random sample $\boldsymbol{x} = \left(x_1 \ldots x_N\right)^T$ is
Equation:

$$\mathrm{f}_{\theta}\left(\boldsymbol{x}\right) = \prod_{n=1}^{N} \mathrm{f}_{\theta}\left(x_n\right) \prod_{n=1}^{N} \theta^k (1-\theta)^{N-k}$$

where $k = \sum_{n=1}^{N} x_n$ is the number of 1's in the sample.

We will show that k is a sufficient statistic for $\boldsymbol{x}$. This will entail showing that the conditional probability mass function $\mathrm{f}_{\theta \mid k}\left(\boldsymbol{x}\right)$ does not depend on θ.

The distribution of the number of ones in N independent Bernoulli trials is binomial:

$$\mathrm{f}_{\theta}\left(k\right) = \binom{N}{k} \theta^k (1-\theta)^{N-k}$$

Next, consider the joint distribution of $\left(\boldsymbol{x}, \sum x_n\right)$. We have

$$\mathrm{f}_\theta\left(\boldsymbol{x}\right) = \mathrm{f}_\theta\left(\boldsymbol{x}, \sum x_n\right)$$

Thus, the conditional probability may be written
Equation:

$$\begin{aligned}
\mathrm{f}_{\theta|k}\left(\boldsymbol{x}\right) &= \frac{\mathrm{f}_\theta(\boldsymbol{x},k)}{\mathrm{f}_\theta(k)} \\
&= \frac{\mathrm{f}_\theta(\boldsymbol{x})}{\mathrm{f}_\theta(k)} \\
&= \frac{\theta^k(1-\theta)^{N-k}}{\binom{N}{k}\theta^k(1-\theta)^{N-k}} \\
&= \frac{1}{\binom{N}{k}}
\end{aligned}$$

This shows that k is indeed a sufficient statistic for θ. The N values $x_1, \ldots, x_N$ can be replaced by the quantity k without losing information about θ.

Exercise:

Problem:

In the previous example, suppose we wish to store in memory the information we possess about θ. Compare the savings, in terms of bits, we gain by storing the sufficient statistic k instead of the full sample $x_1, \ldots, x_N$.

Determining Sufficient Statistics

In the example above, we had to guess the sufficient statistic, and work out the conditional probability by hand. In general, this will be a tedious way to

go about finding sufficient statistics. Fortunately, spotting sufficient statistics can be made easier by the Fisher-Neyman Factorization Theorem.

Uses of Sufficient Statistics

Sufficient statistics have many uses in statistical inference problems. In hypothesis testing, the Likelihood Ratio Test can often be reduced to a sufficient statistic of the data. In parameter estimation, the Minimum Variance Unbiased Estimator of a parameter $\boldsymbol{\theta}$ can be characterized by sufficient statistics and the Rao-Blackwell Theorem.

Minimality and Completeness

Minimal sufficient statistics are, roughly speaking, sufficient statistics that cannot be compressed any more without losing information about the unknown parameter. **Completeness** is a technical characterization of sufficient statistics that allows one to prove minimality. These topics are covered in detail in this module.

Further examples of sufficient statistics may be found in the module on the Fisher-Neyman Factorization Theorem.

Detection in the Presence of Unknowns

We assumed in the previous sections that we have a few well-specified models (hypotheses) for a set of observations. These models were probabilistic; to apply the techniques of statistical hypothesis testing, the models take the form of conditional probability densities. In many interesting circumstances, the exact nature of these densities may not be known. For example, we may know a priori that the mean is either zero or some constant (as in the Gaussian example). However, the variance of the observations may not be known or the value of the non-zero mean may be in doubt. In an array processing context, these respective situations could occur when the background noise level is unknown (a likely possibility in applications) or when the signal amplitude is not known because of far-field range uncertainties (the further the source of propagating energy, the smaller its received energy at each sensor). In an extreme case, we can question the exact nature of the probability densities (everything is not necessarily Gaussian!). The model evaluation problem can still be posed for these situations; we classify the "unknown" aspects of a model testing problem as either **parametric** (the variance is not known, for example) or **nonparametric** (the formula for the density is in doubt). The former situation has a relatively long history compared to the latter; many techniques can be used to approach parametric problems while the latter is a subject of current research (Gibson and Melsa). We concentrate on parametric problems here.

We describe the dependence of the conditional density on a set of parameters by incorporating the parameter vector $\boldsymbol{\theta}$ as part of the condition. We write the likelihood function as $\mathrm{p}_{\boldsymbol{r}\,|\,\mathscr{M}_i\boldsymbol{\theta}}\,(\boldsymbol{r})$ for the parametric problem. In statistics, this situation is said to be a **composite hypothesis** (Cramér). Such situations can be further categorized according to whether the parameters are **random** or **nonrandom**. For a parameter to be random, we have an expression for its a priori density, which could depend on the particular model. As stated many times, a specification of a density usually expresses some knowledge about the range of values a parameter may assume **and** the relative probability of those values. Saying that a parameter has a uniform distribution implies that the values it assumes **are** equally likely, **not** that we have no idea what the values might be and express this

ignorance by a uniform distribution. If we are ignorant of the underlying probability distribution that describes the values of a parameter, we will characterize them simply as being **unknown** (not random). Once we have considered the random parameter case, nonrandom but unknown parameters will be discussed.

Random Parameters

When we know the density of $\boldsymbol{\theta}$, the likelihood function can be expressed as

$$\mathrm{p}_{\boldsymbol{r}|\mathscr{M}_i}\left(\boldsymbol{r}\right) = \int \mathrm{p}_{\boldsymbol{r}|\mathscr{M}_i\theta}\left(\boldsymbol{r}\right) \mathrm{p}_{\boldsymbol{\theta}}\left(\boldsymbol{\theta}\right) \mathrm{d}\,\boldsymbol{\theta}$$

and the likelihood ratio in the random parameter case becomes

$$\Lambda(\boldsymbol{r}) = \frac{\int \mathrm{p}_{\boldsymbol{r}|\mathscr{M}_i\theta}\left(\boldsymbol{r}\right) \mathrm{p}_{\boldsymbol{\theta}}\left(\boldsymbol{\theta}\right) \mathrm{d}\,\boldsymbol{\theta}}{\int \mathrm{p}_{\boldsymbol{r}|\mathscr{M}_i\theta}\left(\boldsymbol{r}\right) \mathrm{p}_{\boldsymbol{\theta}}\left(\boldsymbol{\theta}\right) \mathrm{d}\,\boldsymbol{\theta}}$$

Unfortunately, there are many examples where either the integrals involved are intractable or the sufficient statistic is virtually the same as the likelihood ratio, which can be difficult to compute.

Example:
A simple, but interesting, example that results in a computable answer occurs when the mean of Gaussian random variables is either zero (model 0) or is $\pm(m)$ with equal probability (hypothesis 1). The second hypothesis means that a non-zero mean is present, but its sign is not known. We are therefore stating that if hypothesis one is in fact valid, the mean has fixed sign for each observation - what is random is its a priori value. As before, L statistically independent observations are made.

$$\mathscr{M}_0 : r \sim \mathscr{N}\left(0, \sigma^2 I\right)$$

$$\mathscr{M}_1 : r \sim \mathscr{N}\left(m, \sigma^2 I\right)$$

$$m = \begin{cases} m & \\ \dots & \text{Prob } 1/2 \\ m & \\ -m & \\ \dots & \text{Prob } 1/2 \\ -m & \end{cases}$$

The numerator of the likelihood ratio is the sum of two Gaussian densities weighted by $1/2$ (the a priori probability values), one having a positive mean, the other negative. The likelihood ratio, after simple cancellation of common terms, becomes

$$\Lambda(\boldsymbol{r}) = \frac{1}{2} e^{\frac{2m\sum_{l\ 0}^{L-1} r_l - Lm^2}{2\sigma^2}} + \frac{1}{2} e^{\frac{-2m\sum_{l\ 0}^{L-1} r_l - Lm^2}{2\sigma^2}}$$

and the decision rule takes the form

$$\cosh \quad \frac{m}{\sigma^2} \sum_{l\ 0}^{L-1} r_l \quad \begin{matrix} \mathscr{M}_1 \\ \gtrless \\ \mathscr{M}_0 \end{matrix} \quad \eta e^{\frac{Lm^2}{2\sigma^2}}$$

where $\cosh(x)$ is the **hyperbolic cosine** given simply as $\frac{e^x + e^{-x}}{2}$. As this quantity is an even function, the sign of its argument has no effect on the result. The decision rule can be written more simply as

$$\sum_{l\ 0}^{L-1} r_l \quad \begin{matrix} \mathscr{M}_1 \\ \gtrless \\ \mathscr{M}_0 \end{matrix} \quad \frac{\sigma^2}{|m|} \operatorname{arccosh} \quad \eta e^{\frac{Lm^2}{2\sigma^2}}$$

The sufficient statistic equals the **magnitude** of the sum of the observations in this case. While the right side of this expression, which equals γ, is complicated, it need only be computed once. Calculation of the performance probabilities can be complicated; in this case, the false-alarm probability is easy to find and the others more difficult.

Non-Random Parameters

In those cases where a probability density for the parameters cannot be assigned, the model evaluation problem can be solved in several ways; the methods used depend on the form of the likelihood ratio and the way in which the parameter(s) enter the problem. In the Gaussian problem we have discussed so often, the threshold used in the likelihood ratio test η may be unity. In this case, examination of the resulting computations required reveals that implementing the test **does not require knowledge of the variance of the observations** (see this problem). Thus, if the common variance of the underlying Gaussian distributions is not known, this lack of knowledge has **no effect** on the optimum decision rule. This happy situation - knowledge of the value of a parameter is not required by the optimum decision rule - occurs rarely, but should be checked before using more complicated procedures.

A second fortuitous situation occurs when the sufficient statistic as well as its probability density under one of the models do **not** depend on the unknown parameter(s). Although the sufficient statistic's threshold γ expressed in terms of the likelihood ratio's threshold η depends on the unknown parameters, γ may be computed as a single value using the Neyman-Pearson criterion **if the computation of the false-alarm probability does not involve the unknown parameters**.

Example:
Continuing the example of the previous section, let's consider the situation where the value of the mean of each observation under model $\mathscr{M}_1$ is not known. The sufficient statistic is the sum of the observations (that quantity doesn't depend on m) and the distribution of the observation vector under model $\mathscr{M}_0$ does not depend on m (allowing computation of the false-alarm probability). However, a subtlety emerges; in the derivation of the sufficient statistic, we had to divide by the value of the mean. The critical step occurs once the logarithm of the likelihood ratio is manipulated to obtain

$$m \sum_{l=0}^{L-1} r_l \underset{\mathscr{M}_0}{\overset{\mathscr{M}_1}{\gtrless}} \left(\sigma^2 \ln(\eta) + \frac{Lm^2}{2} \right)$$

Recall that only **positively** monotonic transformations can be applied; if a negatively monotonic operation is applied to this inequality (such as multiplying both sides by -1), the **inequality reverses**. If the sign of m is known, it can be taken into account explicitly and a sufficient statistic results. If, however, the sign is not known, the above expression cannot be manipulated further and the left side constitutes the sufficient statistic for this problem. The sufficient statistic then depends on the unknown parameter and we cannot develop a decision rule in this case. If the sign is known, we can proceed. Assuming the sign of m is positive, the sufficient statistic is the sum of the observations and the threshold γ is found by

$$\gamma = \sqrt{L}\sigma Q^{-1}(P_F)$$

Note that if the variance σ^2 instead of the mean were unknown, we could not compute the threshold. The difficulty lies not with the sufficient statistic (it doesn't depend on the variance), but with the false alarm probability as the expression indicates. Another approach is required to deal with the unknown-variance problem.

When this situation occurs - the sufficient statistic **and** the false-alarm probability can be computed without needing the parameter in question, we have established what is known as a **uniformly most powerful test** (or UMP test) (Cramér; p.529-531), (van Trees; p.89ff). If an UMP test does not exist, which can only be demonstrated by explicitly finding the sufficient statistic and evaluating its probability distribution, then the composite hypothesis testing problem cannot be solved without some value for the parameter being used.

This seemingly impossible situation - we need the value of the parameter that is assumed unknown - can be approached by noting that some data is available for "guessing" the value of the parameter. If a reasonable guess could be obtained, it could then be used in our model evaluation procedures developed in this chapter. **The data available for estimating unknown parameters are precisely the data used in the decision rule**. Procedures intended to yield "good" guesses of the value of a parameter are said to be **parameter estimates**. Estimation procedures are the topic of the next chapter; there we will explore a variety of estimation techniques and develop measure of

estimate quality. For the moment, these issues are secondary; even if we knew the size of the estimation error, for example, the more pertinent issue is how the imprecise parameter value affects the performance probabilities. We can compute these probabilities **without** explicitly determining the estimate's error characteristics.

One parameter estimation procedure that fits nicely into the composite hypothesis testing problem is the **maximum likelihood estimate**. [footnote] Letting $\boldsymbol{r}$ denote the vector of observables and $\boldsymbol{\theta}$ a vector of parameters, the maximum likelihood estimate of $\boldsymbol{\theta}$, $\hat{\theta}_{\mathrm{ML}}$, is that value of $\boldsymbol{\theta}$ that maximizes the conditional density $\mathrm{p}_{\boldsymbol{r}|\boldsymbol{\theta}}\left(\boldsymbol{r}\right)$ of the observations given the parameter values.

To use $\hat{\theta}_{\mathrm{ML}}$ in our decision rule, we estimate the parameter vector **separately** for each model, use the estimated value in the conditional density of the observations, and compute the likelihood ratio. This procedure is termed the **generalized likelihood ratio test** for the unknown parameter problem in hypothesis testing (Lehmann; p.16), (van Trees; p.92ff).
Equation:

$$\Lambda(\boldsymbol{r}) = \frac{\max_{\theta}\left\{\theta, \mathrm{p}_{\boldsymbol{r}|\mathscr{M}_1\theta}\left(\boldsymbol{r}\right)\right\}}{\max_{\theta}\left\{\theta, \mathrm{p}_{\boldsymbol{r}|\mathscr{M}_0\theta}\left(\boldsymbol{r}\right)\right\}}$$

Note that we do **not** find that value of the parameter that (necessarily) maximizes the likelihood ratio. Rather, we estimate the parameter value most consistent with the observed data in the context of each assumed model (hypothesis) of data generation. In this way, the estimate conforms with each potential model rather than being determined by some amalgam of supposedly mutually exclusive models.
The maximum likelihood estimation procedure and its characteristics are fully described in this section.

Example:
Returning to our Gaussian example, assume that the variance σ^2 is known but that the mean under $\mathscr{M}_1$ is unknown.

$$\mathscr{M}_0: \; r \sim \left(0, \sigma^2 I\right)$$

$$\mathscr{M}_1 : \; r \sim \left(m, \sigma^2 I\right)$$

$$m = \begin{matrix} m \\ \dots \\ m \end{matrix}, \; m = ?$$

The unknown quantity occurs only in the exponent of the conditional density under $\mathscr{M}_1$; to maximize this density, we need only to maximize the exponent. Thus, we consider the derivative of the exponent with respect to m.

$$\frac{\partial \left(\left(-\frac{1}{2\sigma^2} \right) \sum_{l=0}^{L-1} \left(r_l - m \right)^2 \right)}{\partial m} \Bigg|_{m=\widehat{m}_{\mathrm{ML}}} = 0 \quad \Rightarrow \quad \sum_{l=0}^{L-1} r_l - \widehat{m}_{\mathrm{ML}} = 0$$

The solution of this equation is the average value of the observations

$$\widehat{m}_{\mathrm{ML}} = \frac{1}{L} \sum_{l=0}^{L-1} r_l$$

To derive the decision rule, we substitute this estimate in the conditional density for $\mathscr{M}_1$. The critical term, the exponent of this density, is manipulated to obtain

$$\left(-\frac{1}{2\sigma^2} \right) \sum_{l=0}^{L-1} \left(r_l - \frac{1}{L} \sum_{k=0}^{L-1} r_k \right)^2 = \left(-\frac{1}{2\sigma^2} \right) \left(\sum_{l=0}^{L-1} r_l^2 - \frac{1}{L} \left(\sum_{l=0}^{L-1} r_l \right)^2 \right)$$

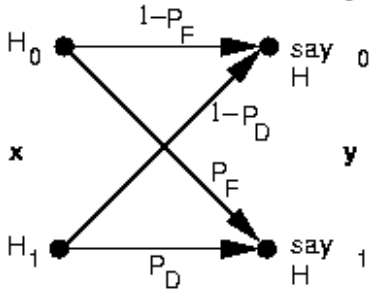
Noting that the first term in this exponent is identical to the exponent of the denominator in the likelihood ratio, the generalized likelihood ratio becomes

$$\Lambda(\boldsymbol{r}) = e^{\frac{1}{2L\sigma^2} \left(\sum_{l=0}^{L-1} r_l \right)^2}$$

The sufficient statistic thus becomes the square (or equivalently the magnitude) of the summed observations. Compare this result with that obtained in [link]. There, an UMP test existed **if** we knew the sign of m and the sufficient statistic was the sum of the observations. Here, where we employed the generalized likelihood ratio test, we made no such assumptions

about m; this generality accounts for the difference in sufficient statistic. Which test do you think would lead to a greater detection probability for a given false-alarm probability?

Once the generalized likelihood ratio is determined, we need to determine the threshold. If the a priori probabilities π_0 and π_1 are known, the evaluation of the threshold proceeds in the usual way. If they are not known, all of the conditional densities must not depend on the unknown parameters lest the performance probabilities also depend upon them. In most cases, the original model evaluation problem is posed in such a way that one of the models does not depend on the unknown parameter; a criterion on the performance probability related to that model can then be established via the Neyman-Pearson procedure. If not the case, the threshold cannot be computed and the threshold must be set experimentally: we force one of the models to be true and modify the threshold on the sufficient statistic until the desired level of performance is reached. Despite this non-mathematical approach, the overall performance of the model evaluation procedure will be optimum because of the results surrounding the Neyman-Pearson criterion.



The two-model testing problem can be abstractly described as a communication channel where the inputs are the models and the outputs are the decisions. The transition

probabilities are related to the false-alarm (P_F) and detection (P_D) probabilities.

Spectral Detection

From the results presented in the previous sections, the colored noise problem was found to be pervasive, but required a computationally difficult detector. The simplest detector structure occurs when the additive noise is white; this notion leads to the idea of whitening the observations, thereby transforming the data into a simpler form (as far as detection theory is concerned). However, the required whitening filter is often time-varying and can have a long-duration unit-sample response. Other, more computationally expedient, approaches to whitening are worth considering. An only slightly more complicated detection problem occurs when we have a diagonal noise covariance matrix, as in the white noise case, but unequal values on the diagonal. In terms of the observations, this situation means that they are contaminated by noise having statistically independent, but unequal variance components: the noise would thus be non-stationary. Few problems fall directly into this category; however, the colored noise problem can be recast into the white, unequal-variance problem by calculating the discrete Fourier Transform (DFT) of the observations and basing the detector on the resulting spectrum. The resulting **spectral detectors** greatly simplify detector structures for discrete-time problems **if** the qualifying assumptions described in sequel hold.

Let W be the so-called $L \times L$ "DFT matrix"

$$W = \begin{matrix} 1 & 1 & 1 & \dots & 1 \\ 1 & W & W^2 & \dots & W^{L-1} \\ 1 & W^2 & W^4 & \dots & W^{2(L-1)} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & W^{L-1} & W^{2(L-1)} & \dots & W^{(L-1)(L-1)} \end{matrix}$$

where W is the elementary complex exponential $e^{-\left(i\frac{2\pi}{L}\right)}$. The discrete Fourier Transform of the sequence $r(l)$, usually written as $R(k) = \sum_{l=0}^{L-1} r(l) e^{-\left(i\frac{2\pi lk}{L}\right)}$, can be written in matrix form as $R = W\boldsymbol{r}$. To analyze the effect of evaluating the DFT of the observations, we describe the computations in matrix form for analytic simplicity. The first critical assumption has been made: take special note that the length of the transform **equals** the duration of the observations. In many signal processing applications, the transform length can differ from the data length, being either longer or shorter. The statistical properties developed in the following discussion are critically sensitive to the equality of these lengths. The covariance matrix K_R of R is given by WK_rW^{H}. Symmetries of these matrices - the Vandermonde form of W and the Hermitian, Toeplitz form of K_r - leads to many simplifications in evaluating this product. The entries no the main diagonal are given by [footnote]

$$K^R{}_{K,K} = \sum_{l=-(L-1)}^{L-1} (L - |l|) K^r_{1,|l|+1} e^{-\left(i\frac{2\pi lk}{L}\right)}$$

The **variance** of the k^{th} term in the discrete Fourier Transform of the noise thus equals the discrete Fourier Transform of the **windowed** covariance function. This window has a triangular shape; colloquially termed the "rooftop" window, its technical name is the **Bartlett window** and

it occurs frequently in array processing and spectral estimation. We have found that the variance equals the smoothed noise power spectrum evaluated at a particular frequency. The off-diagonal terms of K_R are not easily written; the complicated result is

Equation:

$$\forall k_1, k_2, k_1 \neq k_2 : \quad {K^R}_{k_1,k_2} = \sum_{l=0}^{L-1} K^r_{1,l+1} \frac{-1^{k_1-k_2+1} \sin\left(\frac{\pi l (k_1-k_2)}{L}\right)}{\sin\left(\frac{\pi (k_1-k_2)}{L}\right)} \left(e^{i\frac{2\pi l k_1}{L}} + e^{-\left(i\frac{2\pi l k_2}{L}\right)} \right)$$

The complex exponential terms indicate that each off-diagonal term consists of the sum of two Fourier Transforms: one at the frequency index k_2 and the other negative index $-k_1$. In addition, the transform is evaluated only over non-negative lags. The transformed quantity again equals a windowed version of the noise covariance function, but with a **sinusoidal** window whose frequency depends on the indices k_1 and k_2. This window can be negative-valued! In contrast to the Bartlett window encountered in evaluating the on-diagonal terms, the maximum value achieved by the window is not large ($\frac{1}{\sin\left(\frac{\pi (k_1-k_2)}{L}\right)}$ compared to L). Furthermore, this window is **always** zero at the origin, the location of the maximum value of any covariance function. The largest magnitudes of the off-diagonal terms tend to occur when the indices k_1 and k_2 are nearly equal. Let their difference be one; if the covariance function of the noise tends toward zero well within the number of observations, L, then the Bartlett window has little effect on the covariance function while the sinusoidal window greatly reduces it. This condition on the covariance function can be interpreted physically: the noise in this case is wideband and any correlation between noise values does not extend over significant portion of the observation record. On the other hand, if the width of the covariance function is comparable to L, the off-diagonal terms will be significant. This situation occurs when the noise bandwidth is smaller than or comparable to the reciprocal of the observation interval's duration. This condition on the duration of the observation interval relative to the width of the noise correlation function forms the second critical assumption of spectral detection. The off-diagonal terms will thus be much smaller than corresponding terms on the main diagonal $\left({K^R}_{k_1,k_2} \right)^2 \ll {K^R}_{k_1,k_1} {K^R}_{k_2,k_2}$. The curious index $l+1$ on the matrix arises because rows and columns of matrices are traditionally indexed beginning with one instead of zero.

In the simplest case, the covariance matrix of the discrete Fourier Transform of the observations can be well approximated by a diagonal matrix.

$$K_R = \begin{matrix} {\sigma_0}^2 & 0 & \dots & 0 \\ 0 & {\sigma_1}^2 & 0 & \vdots \\ \vdots & 0 & \ddots & 0 \\ 0 & \dots & 0 & {\sigma_{L-1}}^2 \end{matrix}$$

The non-zero components ${\sigma_k}^2$ of this matrix constitute the noise power spectrum at the various frequencies. The signal component of the transformed observations R is represented by S_i, the

DFT of the signal s_i, while the noise component has this diagonal covariance matrix structure.

Note: In the frequency domain, the colored noise problem can be approximately converted to a white noise problem where the components of the noise have unequal variances.

To recap, the critical assumptions of spectral detection are

- The transform length equals that of the observations. In particular, the observations cannot be "padded" to force the transform length to equal a "nice" number (like a power of two).
- The noise's correlation structure should be much less than the duration of the observations. Equivalently, a narrow correlation function means the corresponding power spectrum varies slowly with frequency. If either condition fails to hold, calculating the Fourier Transform of the observations will not necessarily yield a simpler noise covariance matrix.

The optimum spectral detector computes, for each possible signal, the quantity $\Re\left(R^{\mathrm{H}}{K_R}^{-1}S_i\right) - \frac{{S_i}^{\mathrm{H}}{K_R}^{-1}S_i}{2}$ [footnote]. Because of the covariance matrix's simple form, this sufficient statistic for the spectral detection problem has the simple form
Equation:

$$\Re\left(R^{\mathrm{H}}{K_R}^{-1}S_i\right) - \frac{1}{2}{S_i}^{\mathrm{H}}{K_R}^{-1}S_i = \sum_{k=0}^{L-1}\frac{\Re\left(R(k)S_i(k)\right)}{{\sigma_k}^2} - \frac{1}{2}\frac{\left(|S_i(k)|\right)^2}{{\sigma_k}^2}$$

Each term in the dot product between the discrete Fourier Transform of the observations and the signal is weighted by the reciprocal of the noise power spectrum at that frequency. This computation is much simpler than the equivalent time domain version and, because of algorithms such as the fast Fourier Transform, the initial transformation (the multiplication by W or the discrete Fourier Transform) can be evaluated expeditiously.
The real part in the statistic emerges because R and S_i are complex quantities.

Sinusoidal signals are particularly well-suited to the spectral detection approach. **If** the signal's frequency equals one of the analysis frequencies in the Fourier Transform ($\omega_0 = \frac{2\pi k}{L}$ for some k), then the sequence $S_i(k)$ is non-zero only at this frequency index, only one term in the sufficient statistic's summation need be computed, and the noise power is no longer explicitly needed by the detector (it can be merged into the threshold).

$$\Re\left(R^{\mathrm{H}}{K_R}^{-1}S_i\right) - \frac{1}{2}{S_i}^{\mathrm{H}}{K_R}^{-1}S_i = \frac{\Re\left(R(k)S_i(k)\right)}{{\sigma_k}^2} - \frac{1}{2}\frac{\left(|S_i(k)|\right)^2}{{\sigma_k}^2}$$

If the signal's frequency does not correspond to one of the analysis frequencies, spectral energy will be maximal at the nearest analysis frequency but will extend to nearby frequencies also. This effect is termed "leakage" and has been well studied. Exact formulation of the signal's DFT

is usually complicated in this case; approximations which utilize only the maximal-energy frequency component will be sub-optimal (i.e., yield a smaller detection probability). The performance reduction may be small, however, justifying the reduced amount of computation.

Introduction to Estimation Theory

In searching for methods of extracting information from noisy observations, this chapter describes **estimation theory**, which has the goal of **extracting from noise-corrupted observations the values of disturbance parameters (noise variance, for example), signal parameters (amplitude or propagation direction), or signal waveforms**. Estimation theory assumes that the observations contain an information-bearing quantity, thereby tacitly assuming that detection-based preprocessing has been performed (in other words, do I have something in the observations worth estimating?). Conversely, detection theory often requires estimation of unknown parameters: Signal presence is assumed, parameter estimates are incorporated into the detection statistic, and consistency of observations and assumptions tested. Consequently, detection and estimation theory form a symbiotic relationship, each requiring the other to yield high-quality signal processing algorithms.

Despite a wide variety of error criteria and problem frameworks, the optimal detector is characterized by a single result: the likelihood ratio test. Surprisingly, optimal detectors thus derived are usually easy to implement, not often requiring simplification to obtain a feasible realization in hardware or software. In contrast to detection theory, no fundamental result in estimation theory exists to be summoned to attack the problem at hand. The choice of error criterion and its optimization heavily influences the form of the estimation procedure. Because of the variety of criterion-dependent estimators, arguments frequently rage about which of several optimal estimators is "better." Each procedure is optimum for its assumed error criterion; thus, the argument becomes which error criterion best describes some intuitive notion of quality. When more ad hoc, noncriterion-based procedures[footnote] are used, we cannot assess the quality of the resulting estimator relative to the best achievable. As shown later, bounds on the estimation error do exist, but their tightness and applicability to a given situation are always issues in assessing estimator quality. At best, estimation theory is less structured than detection theory. Detection is science, estimation art. Inventiveness coupled with an understanding of the problem (what types of errors are critically important, for example) are key

elements to deciding which estimation procedure "fits" a given problem well.
This governmentese phrase concisely means guessing.

Terminology in Estimation Theory

More so than detection theory, estimation theory relies on jargon to characterize the properties of estimators. Without knowing any estimation technique, let's use parameter estimation as our discussion prototype. The parameter estimation problem is to determine from a set of L observations, represented by the L-dimensional vector $\boldsymbol{r}$, the values of parameters denoted by the vector $\boldsymbol{\theta}$. We write the **estimate** of this parameter vector as $\hat{\theta}(\boldsymbol{r})$, where the "hat" denotes the estimate, and the functional dependence on $\boldsymbol{r}$ explicitly denotes the dependence of the estimate on the observations. This dependence is always present[footnote], but we frequently denote the estimate compactly as $\hat{\theta}$. Because of the probabilistic nature of the problems considered in this chapter, a parameter estimate is itself a random vector, having its own statistical characteristics. The **estimation error** $\varepsilon(\boldsymbol{r})$ equals the estimate minus the actual parameter value: $\varepsilon(\boldsymbol{r}) = \hat{\theta}(\boldsymbol{r}) - \theta$. It too is a random quantity and is often used in the criterion function. For example, the **mean-squared error** is given by $E\left[\boldsymbol{\varepsilon}^T\boldsymbol{\varepsilon}\right]$; the minimum mean-squared error estimate would minimize this quantity. The mean-squared error matrix is $E\left[\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^T\right]$; on the main diagonal, its entries are the mean-squared estimation errors for each component of the parameter vector, whereas the off-diagonal terms express the correlation between the errors. The **mean-squared estimation error** $E\left[\boldsymbol{\varepsilon}^T\boldsymbol{\varepsilon}\right]$ equals the trace of the mean-squared error matrix $\operatorname{tr}\left(E\left[\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^T\right]\right)$.
Estimating the value of a parameter given no data may be an interesting problem in clairvoyance, but not in estimation theory.

Bias

An estimate is said to be **unbiased** if the expected value of the estimate equals the true value of the parameter: $E\left[\hat{\theta}|\theta\right] = \theta$. Otherwise, the estimate is said to be **biased**: $E\left[\hat{\theta}|\theta\right] \neq \theta$. The **bias** $\boldsymbol{b}(\theta)$ is usually considered to be additive, so that $\boldsymbol{b}(\theta) = E\left[\hat{\theta}|\theta\right] - \theta$. When we have a biased estimate, the bias usually depends on the number of observations L. An estimate is said to be **asymptotically unbiased** if the bias tends to zero for large L: $\lim_{L\to\infty} \boldsymbol{b} = 0$. An estimate's variance equals the mean-squared estimation error **only** if the estimate is unbiased.

An unbiased estimate has a probability distribution where the mean equals the actual value of the parameter. Should the lack of bias be considered a desirable property? If many unbiased estimates are computed from statistically independent sets of observations having the same parameter value, the average of these estimates will be close to this value. This property does **not** mean that the estimate has less error than a biased one; there exist biased estimates whose mean-squared errors are smaller than unbiased ones. In such cases, the biased estimate is usually asymptotically unbiased. Lack of bias is good, but that is just one aspect of how we evaluate estimators.

Consistency

We term an estimate **consistent** if the mean-squared estimation error tends to zero as the number of observations becomes large: $\lim_{L\to\infty} E\left[\boldsymbol{\varepsilon}^T\boldsymbol{\varepsilon}\right] = 0$. Thus, a consistent estimate must be at least asymptotically unbiased. Unbiased estimates do exist whose errors never diminish as more data are collected: Their variances remain nonzero no matter how much data are available. Inconsistent estimates may provide reasonable estimates when the amount of data is limited, but have the counterintuitive property that the quality of the estimate does not improve as the number of observations increases. Although appropriate in the proper circumstances (smaller mean-squared error than a consistent estimate over a pertinent range of values of L, consistent estimates are usually favored in practice.

Efficiency

As estimators can be derived in a variety of ways, their error characteristics must always be analyzed and compared. In practice, many problems and the estimators derived for them are sufficiently complicated to render analytic studies of the errors difficult, if not impossible. Instead, numerical simulation and comparison with lower bounds on the estimation error are frequently used instead to assess the estimator performance. An **efficient estimate** has a mean-squared error that equals a particular lower bound: the Cramér-Rao bound. If an efficient estimate exists (the Cramér-Rao bound is the greatest lower bound), it is optimum in the mean-squared sense: No other estimate has a smaller mean-squared error (see Maximum Likelihood Estimators for details).

For many problems no efficient estimate exists. In such cases, the Cramér-Rao bound remains a lower bound, but its value is smaller than that achievable by any estimator. How much smaller is usually not known. However, practitioners frequently use the Cramér-Rao bound in comparisons with numerical error calculations. Another issue is the choice of mean-squared error as the estimation criterion; it may not suffice to pointedly assess estimator performance in a particular problem. Nevertheless, every problem is usually subjected to a Cramér-Rao bound computation and the existence of an efficient estimate considered.

Time-Delay Estimation

An important signal parameter estimation problem is time-delay estimation. Here the unknown is the time origin of the signal: $s(l,\theta) = s(l-\theta)$. The duration of the signal (the domain over which the signal is defined) is assumed brief compared with the observation interval L. Although in continuous time the signal delay is a continuous-valued variable, in discrete time it is not. Consequently, the maximum likelihood estimate **cannot** be found by differentiation, and we must determine the maximum likelihood estimate of signal delay by the most fundamental expression of the maximization procedure. Assuming Gaussian noise, the maximum likelihood estimate of delay is the solution of

$$\min_\theta \left\{\theta, (\boldsymbol{r}-\boldsymbol{s}(\theta))^T {K_n}^{-1} (\boldsymbol{r}-\boldsymbol{s}(\theta))\right\}$$

The term $\boldsymbol{s}^T {K_n}^{-1} \boldsymbol{s}$ is usually assumed not to vary with the presumed time origin of the signal because of the signal's short duration. If the noise is white, this term is constant except near the "edges" of the observation interval. If not white, the kernel of this quadratic form is equivalent to a whitening filter. As discussed later, this filter may be time varying. For noise spectra that are rational and have only poles, the whitening filter's unit-sample response varies only near the edges (see the example). Thus, near the edges, this quadratic form varies with presumed delay and the maximization is analytically difficult. Taking the "easy way out" by ignoring edge effects, the estimate is the solution of

$$\max_\theta \left\{\theta, \boldsymbol{r}^T {K_n}^{-1} \boldsymbol{s}(\theta)\right\}$$

Thus, the delay estimate is the signal time origin that maximizes the matched filter's output.

In addition to the complexity of finding the maximum likelihood estimate, the discrete-valued nature of the parameter also calls into question the use of the Cramér-Rao bound. One of the fundamental assumptions of the bound's derivation is the differentiability of the likelihood function with respect to the parameter. Mathematically, a sequence cannot be differentiated with respect to the integers. A sequence can be differentiated with respect to its argument if we consider the variable to be continuous valued. This approximation can be used only if the sampling interval, unity for the integers, is dense with respect to variations of the sequence. This condition means that the signal must be oversampled to apply the Cramér-Rao bound in a meaningful way. Under these conditions, the mean-squared estimation error for **unbiased estimators** can be no smaller than the Cramér-Rao bound, which is given by

$$E\left[\varepsilon^2\right] \geq \frac{1}{\sum_{k,l\{k,l\}} {K_n}^{-1}{}_{k,l} s'(k-\theta) s'(l-\theta)}$$

which, in the white-noise case, becomes
Equation:

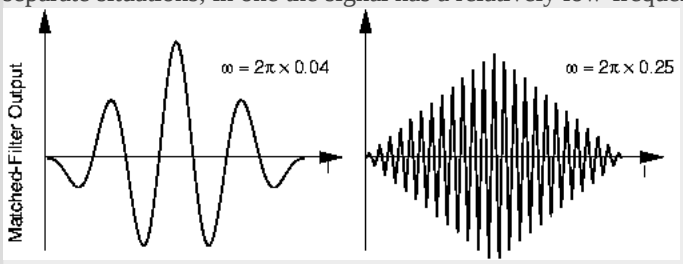
$$E\left[\varepsilon^2\right] \geq \frac{{\sigma_n}^2}{\sum_l \left(s'(l)\right)^2}$$

Here, $s'(\cdot)$ denotes the "derivative" of the discrete-time signal. To justify using this Cramér-Rao bound, we must face the issue of whether an unbiased estimator for time delay **exists**. No general answer exists; each estimator, including the maximum likelihood one, must be examined individually.

Example:
Assume that the noise is white. Because of this assumption, we determine the time delay by maximizing the match-filtered observations.

$$\underset{\theta}{\operatorname{argmax}} \sum_{ll} r(l)s(l-\theta) = \hat{\theta}_{\mathrm{ML}}$$

The number of terms in the sum equals the signal duration. [link] illustrates the match-filtered output in two separate situations; in one the signal has a relatively low-frequency spectrum as compared with the second.



The matched filter outputs are shown for two separate signal situations. In each case, the observation interval (100 samples), the signal's duration (50 samples) and energy (unity) are the same. The difference lies in the signal waveform; both are sinusoids with the first having a frequency of $(2\pi)0.04$ and the second $(2\pi)0.25$. Each output is the signal's autocorrelation function. Few, broad peaks characterize the low-frequency example whereas many narrow peaks are found in the high frequency one.

Because of the symmetry of the autocorrelation function, the estimate **should** be unbiased so long as the autocorrelation function is completely contained within the observation interval. Direct proof of this claim is left to the masochistic reader. For sinusoidal signals of energy E and frequency ω_0, the Cramér-Rao bound is given by $E\left[\varepsilon^2\right] = \frac{{\sigma_n}^2}{{\omega_0}^2 E}$. This bound on the error is accurate only if the measured maximum frequently occurs in the dominant peak of the signal's autocorrelation function. Otherwise, the maximum likelihood estimate "skips" a cycle and produces values concentrated near one of the smaller peaks. The interval between zero crossings of the dominant peak is $\frac{\pi}{2\omega_0}$; the signal-to-noise ratio $\frac{E}{{\sigma_n}^2}$ must exceed $\frac{4}{\pi^2}$ (about 0.5). Remember that this result implicitly assumed a low-frequency sinusoid. The second example demonstrates that cycle skipping occurs more frequently than this guideline suggests when a high-frequency sinusoid is used.

The size of the errors encountered in the time-delay estimation problem can be more accurately assessed by a bounding technique tailored to the problem: the Ziv-Zakai bound (Wiess and Weinstein, Ziv and Zakai). The derivation of this bound relies on results from detection theory (Chazan, Zakai, and Ziv). [footnote] Consider the detection problem in which we must distinguish the signals $s(l-\tau)$ and $s(l-\tau+\Delta)$ while observing them in the presence of white noise that is not necessarily Gaussian. Let hypothesis $\mathscr{M}_0$ represent the case in which the delay, denoted by our parameter symbol θ, is τ and $\mathscr{M}_1$ the case in which $\theta = \tau + \Delta$. The **suboptimum** test statistic consists of estimating the delay, then determining the closest a priori delay to the estimate.

$$\hat{\theta} \underset{\mathscr{M}_0}{\overset{\mathscr{M}_1}{\gtrless}} \tau + \frac{\Delta}{2}$$

By using this ad hoc hypothesis test as an essential part of the derivation, the bound can apply to many situations. Furthermore, by not restricting the type of parameter estimate, the bound applies to any estimator. The probability of error for the optimum hypothesis test (derived from the likelihood ratio) is denoted by $P_e(\tau, \Delta)$. Assuming equally likely hypotheses, the probability of error resulting from the ad hoc test must be greater than that of the optimum.

$$P_e(\tau, \Delta) \leq 1/2 \Pr\left[\varepsilon > \frac{\Delta}{2} \;\middle|\; \mathscr{M}_0\right] + 1/2 \Pr\left[\varepsilon < -\frac{\Delta}{2} \;\middle|\; \mathscr{M}_1\right]$$

Here, ε denotes the estimation error appropriate to the hypothesis.

$$\varepsilon = \begin{cases} \hat{\theta} - \tau \;\text{ if }\; \text{under} \mathscr{M}_0 \\ \hat{\theta} - \tau - \Delta \;\text{ if }\; \text{under} \mathscr{M}_1 \end{cases}$$

The delay is assumed to range uniformly between 0 and L. Combining this restriction to the hypothesized delays yields bounds on both τ and Δ: $0 \leq \tau < L - \Delta$ and $0 \leq \Delta < L$. Simple manipulations show that the integral of this inequality with respect to τ over the possible range of delays is given by [footnote]

$$\int_0^{L-\Delta} P_e(\tau, \Delta)\, \mathrm{d}\, \tau \leq 1/2 \int_0^L \Pr\left[|\varepsilon| > \frac{\Delta}{2} \;\middle|\; \mathscr{M}_0\right] \mathrm{d}\, \tau$$

Note that if we define $\frac{L}{2}\widetilde{P}\left(\frac{\Delta}{2}\right)$ to be the right side of this equation so that

$$\widetilde{P}\left(\frac{\Delta}{2}\right) = \frac{1}{L} \int_0^L \Pr\left[|\varepsilon| > \frac{\Delta}{2} \;\middle|\; \mathscr{M}_0\right] \mathrm{d}\, \tau$$

$\widetilde{P}(\cdot)$ is the complementary distribution function [footnote] of the magnitude of the average estimation error. Multiplying $\widetilde{P}\left(\frac{\Delta}{2}\right)$ by Δ and integrating, the result is

$$\int_0^L \Delta \widetilde{P}\left(\frac{\Delta}{2}\right) \mathrm{d}\, \Delta = -2 \int_0^{\frac{L}{2}} x^2 \frac{\mathrm{d}\, \widetilde{P}}{\mathrm{d}\, x}\, \mathrm{d}\, x$$

The reason for these rather obscure manipulations is now revealed: Because $\widetilde{P}(\cdot)$ is related to the probability distribution function of the absolute error, the right side of this equation is twice the mean-squared error $E\left[\varepsilon^2\right]$. The general Ziv-Zakai bound for the mean-squared estimation error of signal delay is thus expressed as

$$E\left[\varepsilon^2\right] \geq \frac{1}{L} \int_0^L \Delta \int_0^{L-\Delta} P_e(\tau, \Delta)\, \mathrm{d}\, \tau\, \mathrm{d}\, \Delta$$

In many cases, the optimum probability of error $P_e(\tau, \Delta)$ does not depend on τ, the time origin of the observations. This lack of dependence is equivalent to ignoring edge effects and simplifies calculation of the bound. Thus, the Ziv-Zakai bound for time-delay estimation relates the mean-squared estimation error for delay to the probability of error incurred by the optimal detector that is deciding whether a nonzero delay is present or not.
Equation:

$$E\left[\varepsilon^2\right] \geq \frac{1}{L} \int_0^L \Delta\, (L - \Delta) P_e(\Delta)\, \mathrm{d}\, \Delta \geq \frac{L^2}{6} P_e(L) - \int_0^L \left(\frac{\Delta^2}{2} - \frac{\Delta^3}{3L}\right) \frac{\mathrm{d}\, P_e}{\mathrm{d}\, \Delta}\, \mathrm{d}\, \Delta$$

To apply this bound to time-delay estimates (unbiased or not), the optimum probability of error for the type of noise and the relative delay between the two signals must be determined. Substituting this expression into either integral yields the Ziv-Zakai bound.
This result is an example of detection and estimation theory complementing each other to advantage.
Here again, the issue of the discrete nature of the delay becomes a consideration; this step in the derivation implicitly assumes that the delay is continuous valued. This approximation can be greeted more readily as it involves integration rather than differentiation (as in the Cramér-Rao bound).
The complementary distribution function of a probability distribution function $P(x)$ is defined to be $\widetilde{P}(x) = 1 - P(x)$, the probability that a random variable exceeds x.

The general behavior of this bound at parameter extremes can be evaluated in some cases. Note that the Cramér-Rao bound in this problem approaches infinity as either the noise variance grows or the observation interval shrinks to 0 (either forces the signal-to-noise ratio to approach 0). This result is unrealistic as the actual delay is bounded, lying between 0 and L. In this very noisy situation, one should ignore the observations and "guess" **any** reasonable value for the delay; the estimation error is smaller. The probability of error approaches $1/2$ in this situation no matter what the delay Δ may be. Considering the simplified form of the Ziv-Zakai bound, the integral in the second form is 0 in this extreme case.

$$E\left[\varepsilon^2\right] \geq \frac{L^2}{12}$$

The Ziv-Zakai bound is exactly the variance of a random variable uniformly distributed over $[0, L-1]$. The Ziv-Zakai bound thus predicts the size of mean-squared errors more accurately than does the Cramér-Rao bound.

Example:
Let the noise be Gaussian of variance ${\sigma_n}^2$ and the signal have energy E. The probability of error resulting from the likelihood ratio test is given by

$$P_e(\Delta) = Q\left(\sqrt{\frac{E}{2{\sigma_n}^2}\left(1-\rho(\Delta)\right)}\right)$$

The quantity $\rho(\Delta)$ is the normalized autocorrelation function of the signal evaluated at the delay Δ.

$$\rho(\Delta) = \frac{1}{E}\sum_l s(l)s(l-\Delta)$$

Evaluation of the Ziv-Zakai bound for a general signal is very difficult in this Gaussian noise case. Fortunately, the normalized autocorrelation function can be bounded by a relatively simple expression to yield a more manageable expression. The key quantity $1-\rho(\Delta)$ in the probability of error expression can be rewritten using Parseval's Theorem.

$$1-\rho(\Delta) = \frac{1}{2\pi E}\int_0^\pi 2(|S(\omega)|)^2 \times (1-\cos(\omega\Delta))\ \mathrm{d}\ \omega$$

Using the inequality $1-\cos(x) \leq x^2$, $1-\rho(\Delta)$ is bounded from above by $\min\left\{\frac{\Delta^2\beta^2}{2}, 2\right\}$, where β is the root-mean-squared (**RMS**) signal bandwidth.
Equation:

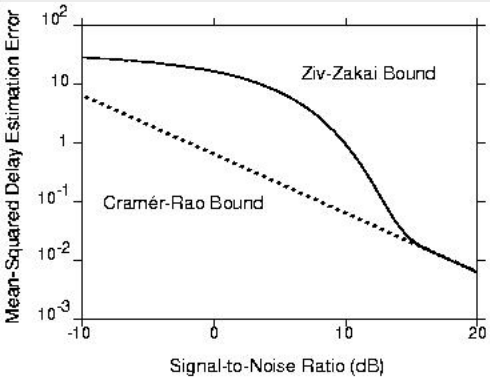
$$\beta^2 = \frac{\int_{-\pi}^{\pi}\omega^2(|S(\omega)|)^2\ \mathrm{d}\ \omega}{\int_{-\pi}^{\pi}(|S(\omega)|)^2\ \mathrm{d}\ \omega}$$

Because $Q(\cdot)$ is a decreasing function, we have $P_e(\Delta) \geq Q\left(\mu\ \min\left\{\Delta, \Delta^*\right\}\right)$, where μ is a combination of all of the constants involved in the argument of $Q(\cdot)$: $\mu = \sqrt{\frac{E\beta^2}{4{\sigma_n}^2}}$. This quantity varies with the product of the signal-to-noise ratio $\frac{E}{{\sigma_n}^2}$ and the squared RMS bandwidth β^2. The parameter $\Delta^* = \frac{2}{\beta}$ is known as the **critical delay** and is twice the reciprocal RMS bandwidth. We can use this lower bound for the probability of error in the Ziv-Zakai bound to produce a lower bound on the mean-squared estimation error. The integral in the first form of the bound yields the complicated, but computable result

$$E\left[\varepsilon^2\right] \geq \frac{L^2}{6}Q\left(\mu\ \min\left\{L, \Delta^*\right\}\right) + \frac{1}{4\mu^2}\ \mathrm{P}_{\chi_3^2}\left(\mu^2\ \min\left\{L^2, {\Delta^*}^2\right\}\right) - \frac{2}{3\sqrt{2\pi}L\mu^3}\left(1-\left(1+\frac{\mu^2}{2}\ \min\left\{L^2\right.\right.\right.$$

The quantity $\mathrm{P}_{\chi_3^2}(\cdot)$ is the probability distribution function of a χ^2 random variable having three degrees of freedom.[footnote] Thus, the threshold effects in this expression for the mean-squared estimation error depend on the relation between the critical delay and the signal duration. In most cases, the minimum equals the critical delay Δ^*, with the opposite choice possible for very low bandwidth signals.

This distribution function has the "closed-form" expression $\mathrm{P}_{\chi_3^2}(x) = \left(1 - Q\left(\sqrt{x}\right) - \sqrt{\frac{x}{2}} e^{-\frac{x}{2}}\right)$.



The Ziv-Zakai bound and the Cramér-Rao bound for the estimation of the time delay of a signal observed in the presence of Gaussian noise is shown as a function of the signal-to-noise ratio. For this plot, $L = 20$ and $\beta = (2\pi)0.2$. The Ziv-Zakai bound is much larger than the Cramér-Rao bound for signal-to-noise ratios less than 13 dB; the Ziv-Zakai bound can be as much as 30 times larger.

The Ziv-Zakai bound and the Cramér-Rao bound for the time-delay estimation problem are shown in [link]. Note how the Ziv-Zakai bound matches the Cramér-Rao bound only for large signal-to-noise ratios, where they both equal $1/4\mu^2 = \frac{{\sigma_n}^2}{E\beta^2}$. For smaller values, the former bound is much larger and provides a better indication of the size of the estimation errors. These errors are because of the "cycle skipping" phenomenon described earlier. The Ziv-Zakai bound describes them well, whereas the Cramér-Rao bound ignores them.

Introduction to Active Sonar

Active sonar processing involves detection, classification and localization. We need to detect if an echo is present, decide if the echo comes from a target or clutter, and determine the position and velocity of the target that formed the echo. These are the fundamental operations of an active sonar system.

During the target echo formation process, other signals are generated by the transmitted sound. Sound will scatter and reflect off of the ocean bottom and surface causing reverberation and direct blast arrivals. Ambient noise, generated by distant shipping and waves breaking on the ocean surface will also be present.

The direct blast corresponds to the sound energy that travels from the source to the receiver via a direct path or with only specular reflections from the surface and bottom. Specular reflection is the mirror like reflection from a surface, and means that the sound ray's angle of incidence and angle of departure are equal. The direct blast energy arrives as a set of discrete arrivals spaced out in time, corresponding to the number of surface and bottom bounces (or reflections) that occurred for each arrival. Each surface and bottom interaction involves some loss, so that arrivals that incur more surface and bottom reflections arrive later in time and are lower in amplitude.

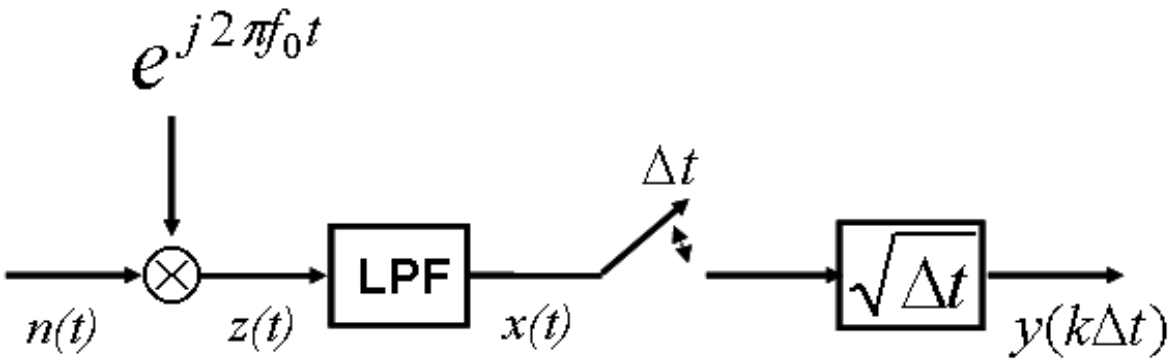
Reverberation corresponds to the diffuse scattering from the surface and bottom. When the sound energy interacts with either boundary, it scatters in all directions, with most of the energy following a specular reflection path. The non-specular scattering continues to propagate to the receiver and generates a continuous, overlapping set of signal replicas termed reverberation. Reverberation has a smooth, continuous power versus time profile. Direct blast arrivals are discrete.

Echoes, direct blast and reverberation will become louder as the transmitted signal gets louder.

Sonar Receiver Model

Active Sonar Receiver Model

Consider the sonar receiver processing chain shown below:



The sonar array input is heterodyned, low-pass filtered, sampled and scaled to generate a discrete time set of samples of the noise plus signal.

The input noise $n(t)$is a real-valued, wide sense stationary random process with power spectral density $P_{\mathrm{nn}}(f)$. Because $n(t)$is wide sense stationary, the power spectral density and the autocorrelation function are related by

$$R_{\mathrm{nn}}(\tau) = \int_{-\infty}^{\infty} P_{\mathrm{nn}}(f) e^{-j2\pi f\tau} \mathrm{df}$$

We assume that $n(t)$is zero mean. A complex carrier $e^{j2\pi \mathrm{ft}}$is applied to the random process $n(t)$resulting in a complex random process $z(t)$. The mean value of $z(t)$is given by:

$$E\{z(t)\} = E\{n(t)e^{j2\pi \mathrm{ft}}\} = E\{n(t)\}e^{j2\pi \mathrm{ft}} = 0$$

The covariance of $z(t)$is given by

$$E\{z(t)z(s)\} = e^{j2\pi f(t-s)} E\{n(t)n(s)\} = e^{j2\pi f(t-s)} R_{\mathrm{nn}}(t-s),$$

which shows that $z(t)$is wide sense stationary as well.

$z(t)$is passed through a band-pass filter to produce $x(t)$. The frequency response of the band-pass filter is assumed to be low-pass with a bandwidth of $\mathrm{B}/2$Hertz. That is we will assume that the filter transfer function $H_{\mathrm{BPF}}(f)$is given by:

$$H(f) = \begin{cases} 1, \mid f \mid < B/2 \\ 0, \mid f \mid > B/2 \end{cases}$$

The resulting $x(t)$is a wide sense stationary random process with zero-mean and power spectral density:

$$P_{\mathrm{xx}}(f) \approx \begin{cases} N_0/2, \mid f \mid < B/2 \\ 0, \mid f \mid > B/2 \end{cases}, \quad (1)$$

Where we have assumed that the bandwidth of the receiver is small relative to the center frequency of the signal we are trying to detect, $B/f_0 << 1$. The power spectral density of $x(t)$can then be approximated by the power spectral density of the noise near f_0:

$$P_{\mathrm{nn}}(f) \approx P_{\mathrm{nn}}(f_0) = N_0/2, \mid f - f_0 \mid < B/2$$

If $x(t)$ has a power spectral density given by Eq-1, then the autocorrelation function of $x(t)$ becomes:

$$R_{\mathrm{xx}}(\tau) = E\{x(t)x(t+\tau)\} = \int_{-B/2}^{B/2} \frac{N_0}{2} e^{j2\pi f\tau} \mathrm{df} = \frac{N_0}{2} \frac{\sin \pi B\tau}{\pi \mathrm{\tau}}$$

Note that $R_{\mathrm{xx}}(0) = \frac{N_0 B}{2}$.

Now if we choose a sampling interval $\Delta t = 1/B$; then the samples at $k\Delta t$ have an autocorrelation given by

$$E\{x(k\Delta t)x(l\Delta t)\} = \frac{N_0}{2} \frac{\sin(\pi B(k-l)\Delta t)}{\pi(k-l)\Delta t} = \frac{\mathrm{BN}_0}{2} \delta_{\mathrm{kl}},$$

Hence $x(k\Delta t)$,$k = 0$,1,... is a discrete time, wide sense stationary, white noise with intensity $\frac{\mathrm{BN}_0}{2}$.

For matched filtering applications, we scale the output of the Analog to Digital conversion process by $\Delta t = 1/B$to conserve the signal energy over a time interval T. This creates the discrete time process $y_k = x(k\Delta t)\sqrt{\Delta t}$.

To see this, consider that

$$E\left\{\int_0^T \mid x(t) \mid^2 \mathrm{dt}\right\} = \int_0^T E\{\mid x(t) \mid^2\}\mathrm{dt} = \int_0^T R_{\mathrm{xx}}(0)\mathrm{dt} = \frac{\mathrm{BTN}_0}{2}$$

And

$$E\left\{\sum_{k=1}^{k=\frac{T}{\Delta t}} \mid y(k) \mid^2\right\} = \sum_{k=1}^{k=\frac{T}{\Delta t}} E\left\{\mid y(k) \mid^2\right\} = \sum_{k=1}^{k=\frac{T}{\Delta t}} \frac{\mathrm{BN}_0}{2}\Delta t = \frac{\mathrm{BTN}_0}{2}$$

.

Active Sonar Detection in Ambient Noise

Introduction

In a sonar system, one is often searching for targets, e.g. submarines, mines, or fish. The sonar system gathers sounds from its acoustic sensors searching for either echoes or sounds emitted by the target. These received sounds are the observations, and for active sonar, are collected into sets of observations related to the time of transmission of the sonar waveform. The sounds related to the broadcast of a single transmission is called a ping history. These pings occur sequentially in time, so one naturally has a sequence of observations (sound recordings) indexed by the time that a ping was transmitted.

The sonar system decision space includes hypotheses about the target's presence, location, velocity and classification. The observations Y_k at ping k contain information about the sonar decision space, but are also influenced, and often dominated by other sounds, such as noise and reverberation. Echoes, noise and reverberation are significantly influenced by the propagation properties of the ocean. These environmental effects are important when making useful inferences about the target echoes that may or may not be present in the sonar ping history.

In most sonar systems today, environmental interference effects, such as noise and reverberation, are treated as random variables. The sonar processing designer develops algorithms that make detections and estimates target states by assuming a statistical model of the echo and interference, choosing environmental interference model parameters (amplitude, covariance, autocorrelation, etc.) and then computing a detection decision or state estimate.

The environmental effects are usually estimated as part of the target state decision process, or the processing algorithm is constructed to be invariant to the environmental effects.

The primary detection processing method for current active sonars is to process the ping history with a bank of matched filters. The filters are constructed so that each filter is constructing the cross-correlation between the transmitted waveform and the pre-whitened ping history.

A monostatic sonar has the source and receiver in the same location, and hence the receiver cannot realistically capture the ping history until the waveform transmission is complete. To simplify the problem, we will look for targets that are far away from the sonar, so that the echo reception occurs when the reverberation level has fallen below the background noise level. In this way, we are dealing with target echoes that are essentially embedded in background noise only.

Decision Space

We use the symbol φ to designate the target absent hypothesis. The other hypotheses concern the location of a single target. We then use as a decision space the composite space $D = \varphi \cup H^{\tau}$where $H^{\tau}$is the space constructed from all target present hypotheses. The target hypothesis space consists of those locations around the sonar that generate echoes embedded in noise only. These location hypothesis h are part of the decision space $h = (x,y) \in H^{\tau}$specified by the set of ordered pairs $h = (x,y)$ such that

$h = (x,y) \in H^{\tau}$if $R_{\min} < \sqrt{(x - x_r)^2 + (y - y_r)^2} < R_{\max}$

Where we are assuming that the depth of the target is small when compared to its (x,y) coordinates, the receiver is located at (x_r,y_r). $R_{\min}$is the range at which the echo is noise, not reverberation limited, and $R_{\max}$is the farthest range of interest. For this problem, h is an index into the target range from the sonar.

Sonar Receiver Model

The sonar transmits the waveform $m(t)$for each ping. In most sonar transmitters, the transmitted waveform is narrow-band, that is, the waveform bandwidth is much smaller than its center frequency, f. This is true because efficient sonar transmitters use resonant mechanical and electrical components to provide maximum electrical to sound power transfer. An approximation therefore is to model the transmitted waveform as an amplitude modulated carrier:

$m(t) = \sin(2\pi \mathrm{ft}) w(t)$, $t = (0,T)$

We will assume that the target is motionless, so that Doppler effects can be ignored. We will assume that the sonar receiver is a single sensor, with no directionality characteristics. For each target location hypothesis $h = (x,y)$ we know approximately the received echo time series:

$g(t \mid h) = \mathrm{Bm}(t - 2\mathrm{R}/c)$

The amplitude B is related to the propagation loss out to the target hypothesis location, and the reflection characteristics of the target. The time delay $2\mathrm{R}/c$ corresponds to the time it takes for the transmission waveform to reach the target and return to the sonar. R is the range to the target and c is the effective speed of sound, when including refraction and boundary reflections.

The received echo is band-limited to approximately the same frequency band as the transmission. The receiver bandwidth may be greater than the transmitted bandwidth due to Doppler frequency shifts, but for the present, we are assuming that the target is not moving. Sonar receivers use heterodyne techniques to reduce the data storage of the ping history. The sonar receiver multiplies the ping history by a carrier signal $e^{-j2\pi \mathrm{ft}}$to shift the positive frequency part of the received echo closer to DC. The resulting signal is then low pass filtered to eliminate the shifted negative frequency part of the ping history. Since the original ping history was real, the negative frequency part of the signal spectra carries no additional information. The result is a complex signal with a lower bandwidth, but retains all of the echo related information of the original ping history. This heterodyne process can be done in the analog or digital domain.

A target echo passing through the heterodyne part of the sonar receiver becomes:

$$r(t \mid h) = \mathrm{A}e^{j\theta} w(t - 2\mathrm{R}/c)$$

The phase shift θ corresponds to the phase shift due to heterodyne operation; the uncertainty in propagation conditions; and the summation of multi-path arrivals with almost the same time delay, etc.

We will assume that the target echo amplitude, $\mathrm{A}e^{j\theta}$,is a complex Gaussian random variable with zero mean and with standard deviation $\sigma^2(h)$.We are modeling the echo as having the same waveform as the transmission, but with an uncertain phase and amplitude. This is assuming that the target echo amplitude obeys Swerling target type I statistics with unknown phase.

While receiving an echo, we will also receive ambient noise, $q(t)$, which we will assume to be complex Gaussian noise, with constant power spectral density over the receiver's bandwidth. The noise power spectral density over the receiver's bandwidth BW is assumed to be N_0Pascals^2/Hertz.

We receive L complex valued samples, y as the ping history after heterodyning. For hypothesis h, the observation in discrete time is:

$$y(k\Delta t) = q(k\Delta t) + r(k\Delta t - 2\mathrm{R}/c \mid h) \text{ for }, k = 1, \ldots L$$

where Δt is the digital sample rate after heterodyning, and q is a sample of the noise and reverberation. Note that $r(k\Delta t - 2\mathrm{R}(h)/c \mid h) = 0$ when $2\mathrm{R}(h)/c > k\Delta t$, because the echo is delayed. The delay for hypothesis h in samples, is given by

$$D(h) = \left[\frac{2\mathrm{R}}{c\Delta t}\right]$$

where [x] is the nearest integer to x. We choose the sample rate Δt to be small enough to satisfy the Nyquist sampling criteria for the received echo. We will assume that the non-zero part of the echo is N samples long.

Statistical Model of the Ping History

We will represent the sampled echo response as a partitioned vector:

$$r(h) = \begin{matrix} 0_{D(h)\,x1} \\ Aw \\ 0_{L-D(h)-N} \end{matrix},$$

Where

$$w = \begin{matrix} w(\Delta t) \\ \vdots \\ w(N\Delta t) \end{matrix}$$

and the sampled noise and interference as a vector

$$q = \begin{matrix} q_1 \\ \vdots \\ q_L \end{matrix},$$

so that the sampled ping history becomes

$$y = q + r(h)$$

The echo is modeled as a known signal w, with Gaussian random complex amplitude A, with zero mean and variance $\sigma^2_{A(h)}$. We will assume that $w^H w = 1$, and that $|\ A(h)\ |^2$ is the energy of the echo, with units Pascals^2-seconds. Since $\sigma^2_{A(h)}$is $E|\ A(h)\ |^2$, it has units of Pascals^2-seconds as well. The amplitude of the echo is a function of the target location hypothesis h. The location of $w$in $r(h)$depends on the location of the target through the time delay $D(h)$.

Since each element of the random vector Aw is complex Gaussian, the random vector Awhas a complex Gaussian distribution. The probability density of Aw is Gaussian zero mean with covariance matrix $\sigma^2_{A(h)}$ww^H. To see this, consider that

$E(Aw) = E(A)w = 0_N$

The covariance of Aw is given by:

$E(Aw)(Aw)^H = E(\mathrm{AA}^H)\mathrm{ww}^H = \sigma^2_{A(h)}\mathrm{ww}^H$

hence $r(h)$is zero mean complex Gaussian with covariance matrix $\sigma^2_{A(h)}\mathrm{rr}^H$.

For the clutter only hypothesis φ,$y = q$.

We have sampled, heterodyned and possibly re-sampled the noise process $q(t)$to form q .

During the period where r is non-zero, *q*is a sampled version of the ambient noise, represented as a N by 1 complex Gaussian noise random vector with zero mean and covariance matrix $(N_0)I_N$. This is true because $\mathrm{BW}\Delta t \approx 1$ for complex Nyquist sampling of a band-limited signal.

Overall, the noise and reverberation *q* is assumed to be complex Gaussian with zero mean and L by L covariance matrix C.

Because we are assuming that the reverberation dies away before the echoes from the target search arrive, C has the following partition:

$$C = \begin{bmatrix} R & 0 \\ 0 & N_0 I \end{bmatrix}$$

Matrix R has dimensions of $D_{\min}\mathrm{xD}_{\min}$, the minimum delay where the echo interference is dominated by Ambient noise.

Under target hypothesis h , y is Gaussian with has zero mean and covariance matrix $C + \sigma_{A^2}\mathrm{rr}^H$.

The probability density of $y$under h becomes:

$p(y \mid h) = \frac{1}{\pi^N \det(C_r + C)} \exp\left(-y^H (C_r + C)^{-1} y\right),$

where $C_r = \sigma_{A^2}\mathrm{rr}^H$.

Under the clutter hypothesis, φ, y has zero mean and covariance matrix C.The probability density of $y$under φ becomes:

$p(y \mid \varphi) = \frac{1}{\pi^N \det(C)} \exp\left(-y^H (C)^{-1} y\right)$

Active Sonar Likelihood Function

We will find that the natural logarithm of the measurement likelihood ratio simplifies the detection expression considerably:

$$\log\frac{L(y|h)}{L(y|\varphi)} = \log(\det(C)) - \log(\det(C+C_r)) + y^H\left(C^{-1} - (C_r+C)^{-1}\right)y$$

The determinant of $C$is $\det(C) = \det(N_0 I_{L-D_{\min}})\det(R) = (N_0)^{L-D_{\min}}\det(R)$

It is convenient to partition $C + C_r$into sub-matrices compatible with r.

$$C + C_r = \begin{matrix} R & 0 & 0 & 0 \\ 0 & N_0 I_{D(h)-D_{\min}} & 0 & 0 \\ 0 & 0 & N_0 I_N + \sigma^2_{A(h)}\mathrm{ww}^H & 0 \\ 0 & 0 & 0 & N_0 I_{L-D(h)-N} \end{matrix}$$

The determinant of $C + C_r$is

$$\det(C+C_r) = \det(R)\det(N_0 I_{D(h)-D_{\min}})\det(N_0 I_N + \sigma^2_{A(h)}\mathrm{rr}^H)\det(N_0 I_{L-D(h)-N})$$

Which becomes

$$\det(C+C_r) = (N_0)^{L-D_{\min}-N}\det(N_0 I_N + \sigma^2_{A(h)}\mathrm{rr}^H)\det(R)$$

Using Sylvester's determinant theorem:

$$\det(I + \mathrm{AB}) = \det(I + \mathrm{BA})$$

We obtain:

$$\det(C+C_r) = (N_0)^{L-D_{\min}}(1 + \frac{\sigma^2_{A(h)}r^H r}{N_0})\det(R)$$

Which equals

$$\det(C+C_r) = (N_0)^{L-D_{\min}}(1 + \frac{\sigma^2_{A(h)}}{N_0})\det(R)$$

Now we partition the observed ping history y into

$$y = \begin{matrix} y_R \\ y_{N1} \\ y_h \\ y_{N2} \end{matrix}$$

so that:

$$C^{-1} - (C_r + C)^{-1} = \begin{matrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{N_0} I_N - (N_0 I_N + \sigma^2_{A(h)} \mathrm{w}\mathrm{w}^H)^{-1} & \\ 0 & 0 & 0 & 0 \end{matrix}$$

The Woodbury matrix identity states:

$$(A + \mathrm{UCV})^{-1} = A^{-1} - A^{-1}U(C^{-1} + \mathrm{V}\mathrm{A}^{-1}U)^{-1}\mathrm{V}\mathrm{A}^{-1}$$

So that

$$(N_0 I_N + \sigma^2_{A(h)} \mathrm{w}\mathrm{w}^H)^{-1} = \frac{1}{N_0} I_N - \frac{1}{N_0^2} \frac{\mathrm{w}\mathrm{w}^H}{\frac{1}{N_0} + \frac{1}{\sigma^2_{A(h)}}}$$

And hence

$$y^H \left(C^{-1} - (C_r + C)^{-1}\right) y = \frac{1}{N_0} \frac{\sigma^2_{A(h)}/N_0}{(1 + \frac{\sigma^2_{A(h)}}{N_0})} \mid w^H y_h \mid^2$$

So that the log likelihood function becomes

$$\log \frac{L(y|h)}{L(y|\varphi)} = -\log(1 + \frac{\sigma^2_{A(h)}}{N_0}) + \frac{1}{N_0} \frac{\sigma^2_{A(h)}/N_0}{(1 + \frac{\sigma^2_{A(h)}}{N_0})} \mid w^H y_h \mid^2$$

Putting the noise variance inside the observation vector yields for the log likelihood function

$$\log \frac{L(y|h)}{L(y|\varphi)} = -\log(1 + \frac{\sigma^2_{A(h)}}{N_0}) + \frac{\sigma^2_{A(h)}/N_0}{(1 + \frac{\sigma^2_{A(h)}}{N_0})} \mid w^H \frac{y_h}{\sqrt{N_0}} \mid^2$$

The log-likelihood ratio shows that the assumed echo shape, w^H, the assumed energy of the signal $\sigma^2_{A(h)}$, the ambient noise density N_0, and the expected location in time of the echo needed to select $y_h$are the parameters required to evaluate the log-likelihood

function. The assumed echo shape and energy can vary by hypothesis h, but the noise properties N_0have been assumed to be the same for each decision hypothesis h.

Signal to Noise Ratio of the Detector

The magnitude squared term, $\mid w^H \frac{y_h}{\sqrt{N_0}} \mid^2$, is a matched filter, where the observations are cross-correlated with the signal template. The observations are normalized by the noise variance before cross correlation, which is a form of pre-whitening. The term $\frac{\sigma^2_{A(h)}}{N_0}$ is the Energy to Noise Density Ratio (ENR) of the detection problem. Note that the signal to noise ratio of the matched filter output can be written as:

$$\frac{E\left\{|w^H \frac{Aw}{\sqrt{N_0}}|^2\right\}}{E\left\{|w^H \frac{q}{\sqrt{N_0}}|^2\right\}} = \frac{\frac{\sigma^2_{A(h)}}{N_0}}{E\left\{\frac{w^H \mathrm{qq}^H w}{N_0}\right\}} = \frac{\frac{\sigma^2_{A(h)}}{N_0}}{w^H \mathrm{I} w} = \frac{\sigma^2_{A(h)}}{N_0} = \mathrm{ENR}$$

This result is a general result for matched filters. A matched filter's SNR is the Energy to Noise Density Ratio of the problem. The energy of a signal being detected is related to the average amplitude and the duration of the signal. The SNR output of the matched filter is independent of the details of the waveform being detected, only the signal energy and the noise spectral density determine the matched filter response.

The likelihood function is dependent only the ENR as well.

A-priori Assumptions

Before the ping history is received, we assess the probability of each hypothesis, $p(h)$ and $p(\varphi)$. The a-priori information may have come from previous pings, or are probabilities assigned by the sonar system to begin a target search.

Using the logarithm of probability density function ratios simplifies the expressions:

$$\log \frac{p(h|y)}{p(\varphi|y)} = \log \frac{L(y|h)}{L(y|\varphi)} + \log \frac{p^-(h)}{p^-(\varphi)}$$

Using the likelihood ratio notation,

$$\log \Lambda(h \mid y) = \log \frac{L(y|h)}{L(y|\varphi)} + \log \Lambda^-(h)$$

Once we compute $\log \Lambda(h \mid y)$, we can declare a target is present, with confidence p_T by computing:

Target Present if: $\int\limits_{h \in H^{\tau}} \Lambda(h) \mathrm{dh} > \frac{1-p_T}{p_T}$

Because the target hypothesis space contains many hypotheses, this detection problem is can be considered a composite hypothesis test.

An alternative approach to detection of a target with an unknown range is solved by finding the target range hypothesis with the greatest measurement likelihood as the detection statistic [Kay]. This is referred to as Generalized Maximum Likelihood Ratio Testing (GLRT).

Target Present if: $\max\limits_{h} \log \frac{L(y|h)}{L(y|\varphi)} > \gamma$

The GLRT approach is often easier to implement than the Bayes detection approach, because one avoids the integration/summation over a-priori probabilities.

Properties of Active Sonar Matched Filtering

Properties of Active Sonar Matched Filtering

Introduction

Matched filters are used extensively in coherent active sonar. The output of a matched filter is used for detection, classification and localization. This document develops some properties of matched filters, including the SNR response in ambient noise and the response to reverberation.

In a matched filter for active sonar, we are integrating the echo plus interference times the echo's replica. When an echo passes through the matched filter, we are cross-correlating the echo with a scaled version of the echo, so that the output is a scaled version of the auto-correlation of the echo corrupted by noise. The autocorrelation of the echo has a peak in time whose duration is approximately the inverse of the echo's bandwidth.

For some waveforms (such as the Sinusoidal Frequency Modulation pulse) the autocorrelation function will have multiple peaks, termed 'fingers', due to the periodic structure of the pulse. Each autocorrelation finger has a time width approximately equal to the signal's bandwidth.

Continuous Time Matched Filter

The echo is written as

$e(t) = \sqrt{E_R} r(t)$, where $\int_0^T r^2(t) \mathrm{dt} = 1$

This implies that the echo energy $\int_0^T e^2(t) \mathrm{dt}$is E_R, measured in Pascal^2-seconds.

We can write the matched filter operation in continuous time as

$$m(t) = \int_t^{t+T} y(\sigma) r(\sigma - t) d\sigma$$

$y(\sigma)$is the receiver time series. In response to a target echo that arrives at $T_D$seconds and without noise or reverberation, the receiver output is $y(t) = e(t - T_D)$. The output of the matched filter becomes:

$$m(t) = \int_t^{t+T} e(\sigma - T_D) r(\sigma - t) d\sigma = \sqrt{E_R} \int_t^{t+T} r(\sigma - T_D) r(\sigma - t) d\sigma$$

Hence $m(T_D) = \sqrt{E_R}$. The peak power output of the matched filter, $m^2(t)$, in response to a echo is E_R.

We determine the matched filter response to noise next. Assume the input noise is white with variance AN_0:

$$E\{n(t) n(s)\} = \mathrm{AN}_0 \delta(t - s)$$

Note that the delta function has units of inverse seconds, and therefore AN_0has units of Pascals^2/Hz, equivalent to a spectral density. From the definition of stationary random process autocorrelations and power spectral density, we know that the Fourier transform of the autocorrelation is the spectral density function for the random process. The Fourier transform of covariance becomes
$\int e^{j2\pi f\tau}\mathrm{AN}_0\delta(\tau)d\tau = \mathrm{AN}_0$, which is the spectral density of the noise.

$$E\{m(t)m(s)\} = E\left\{\int_t^{t+T} n(\sigma)r(\sigma-t)d\sigma \int_t^{t+T} n(\beta)r(\beta-t)d\beta\right\} =$$

$$\mathrm{AN}_0\int_t^{t+T}\int_t^{t+T}\delta(\sigma-\beta)r(\sigma-t)r(\beta-t)d\sigma d\beta = \mathrm{AN}_0\int_t^{t+T} r^2(\sigma-t)d\sigma = \mathrm{AN}_0$$

Thus, the noise power response of a matched filter is the input spectral density, AN_0.

We conclude that the signal to noise ratio (SNR) at the output of a matched filter is the ratio of the echo energy to the noise spectral density, E_R/AN_0. This assumes that the noise is white, e.g. a flat spectral density at the input to the matched filter. This is a general result, independent of the signal waveform details, except for its energy E_R.

Discrete Time Matched Filters

Discrete time filters have nearly the same properties as continuous time filters. In discrete time, we assume an echo of $e(k) = \sqrt{E_R}r(k)$, with $\sum_{k=1}^{T} r^2(k) = 1$. The discrete matched filter output to the input y(k) is given by:

$m(k) = \sum_{l=k}^{k+T-1} y(l)r(l-k+1)$

In response to the echo, $y(k) = e(k-T_D)$the output of the discrete time matched filter is

$m(k) = \sum_{l=k}^{k+T-1} e(l-T_D)r(l-k+1) = \sqrt{E_R}\sum_{l=k}^{k+T-1} r(l-T_D)r(l-k+1)$

Hence $m(T_D-1) = \sqrt{E_R}$. The peak power output of the matched filter, $m^2(t)$, in response to a echo is E_R.

We determine the discrete matched filter response to noise next. Assume the input noise is sampled white with variance AN_0:

$E\{n(k)n(l)\} = \mathrm{AN}_0\delta_{\mathrm{kl}}$

$$E\{m(k)m(p)\} = E\left\{\sum_{l=k}^{k+T-1} n(l)r(l-k+1)\sum_{i=p}^{p+T-1} n(i)r(i-p+1)\right\} =$$

$$\mathrm{AN}_0\sum_{l=k}^{k+T-1}\sum_{i=p}^{p+T-1}\delta_{\mathrm{li}}r(l-k+1)r(i-p+1) = \mathrm{AN}_0\sum_{l=k}^{k+T-1} r^2(l-k+1) = \mathrm{AN}_0$$

Thus, the signal to noise ratio at the output of a discrete time matched filter is E_R/AN_0.

The matched filter compresses the echo signal to a pulse (or a series of pulses for waveforms such as SFM) with time width equal approximately to its inverse bandwidth, 1/BW.

Matched Filter Response to Reverberation

One model for reverberation assumes that the reverberation comes from distributed discrete scatterers, with density $A(u)$.

$$y(t) = \sqrt{E_T} \int_0^\infty A(u)\Gamma(u)r(t-\tau(u))du,$$

$A(u)$ is considered a random, spatial process that models the amplitude of the scattering that occurs at range u back to the receiver. We are assuming that the receiver has significant aperture, and that y(t) is the receiver response at the output of a beamformer. In this case, scattering is occurring from the patch of the ocean bottom or surface that lies at range u and within the receiver beamwidth in azimuth and elevation. Each patch of the bottom or surface will arrive at the receiver at a different time. $\Gamma(u)$is the transmission loss from the source to the scattering range (u) and back to the receiver. $\tau(u)$is the total travel time from source to scatterer to receiver. As one can see, the reverberation is made up of many time delayed and amplitude scaled replicas of the transmitted waveform.

The matched filter response to the reverberation is

$$m_R(t) = \int_t^{t+T} r(\sigma - t)\sqrt{E_T} \int_0^\infty A(u)\Gamma(u)r(\sigma-\tau(u))dud\sigma =$$

$$\sqrt{E_T} \int_0^\infty A(u)\Gamma(u) \int_t^{t+T} r(\sigma-\tau(u))r(\sigma-t)d\sigma du$$

We define the transmitted waveform autocorrelation function as

$$\chi(\tau) = \int_0^T r(t-\tau)r(t)\mathrm{dt}$$

Recall, that by definition, $\chi(0) = \int_0^T r^2(t)\mathrm{dt} = 1$. In more general terms, we define the transmitted wideband signal ambiguity function as

$$\chi_{\mathrm{WB}}(\tau,\eta) = \sqrt{\eta}\int_{-\infty}^{\infty} r\left(\eta(t-\tau)\right)r(t)\mathrm{dt}$$

Note: Some authors define the ambiguity function as the magnitude squared value of this definition. Other authors choose different normalizations or the sign (-/+) on the delay term τ.

In the wideband signal ambiguity function, the Doppler effect is represented by the scaling factor η. In narrowband cases, the Doppler effect is represented by a frequency shift, φ. For a monostatic sonar the frequency shift is given by $\varphi = 2\mathrm{v}/\mathrm{c}$, where $v$is the radial velocity between the scattering object and the sonar system.

$$\chi_{\mathrm{NB}}(\tau,\varphi) = \int_{-\infty}^{\infty} r\left(t-\tau\right)r(t)e^{-j2\pi\varphi t}\mathrm{dt}$$

One can show (Weiss) that the narrowband approximation to Doppler is valid if $2\mathrm{v}/\mathrm{c} << \frac{1}{\mathrm{BT}}$, where $B$is the waveform bandwidth and $T$is the duration. For one hundred (100) Hertz bandwidth waveforms that last for one (1) second, the speed of the target must be much less than 7.5 m/sec, or approximately 15 knots.

An important invariance property of the narrowband ambiguity function is that

$$\int | \chi_{\mathrm{NB}}(\tau,\varphi) |^2 \mathrm{dt} = 1$$

Using either definition of the signal ambiguity function we have

$$\int_t^{t+T} r(\sigma - \tau(u)) r(\sigma - t) d\sigma = \int_0^T r(\sigma' - (\tau(u) - t)) r(\sigma') d\sigma' = \chi(\tau(u) - t,0)$$

Therefore the matched filter response is

$$m_R(t) = \sqrt{E_T} \int_0^\infty A(u) \Gamma(u) \chi(\tau(u) - t,0) du$$

This expression for the matched filter response shows that the amount of reverberation at time t is directly related to the transmitted signal's autocorrelation and energy source level (ESL). Wide band signals will have narrower autocorrelation peaks, and thus less reverberation amplitude.

Using the narrowband approximation for Doppler shifts allows efficient implementation of matched filter banks as generalized spectrogram analysis. One treats the replica as a "window function" in place of the more traditional Hanning or Hamming windows. The matched filter for narrowband Doppler shifts is given by:

$$m(t,\varphi) = \int_t^{t+T} y(\sigma) r(\sigma - t) e^{-j2\pi\varphi\sigma} d\sigma$$

Which can be rewritten as:

$$m(t,\varphi) = e^{j2\pi\varphi t} \int_{-\infty}^{\infty} y(t + \sigma') r(\sigma') e^{-j2\pi\varphi \{\sigma'} d\sigma'$$

The spectrogram with window function $w(\sigma)$ is given by:

$$\int_{-\infty}^{\infty} y(t + \sigma') w(\sigma') e^{-j2\pi\varphi \{\sigma'} d\sigma'$$

$$\Phi(t,\varphi) = ||^2$$

Hence, the squared envelope of the narrowband Doppler matched filter is a spectrogram with the window function being the conjugate of the transmitted waveform replica. This interpretation of narrowband matched filtering lends insight into the use of different window functions on transmitted waveforms. Often one adds a Hanning or Tukey window to the transmitted waveform. This windowing is necessary in some cases because the sonar transmitter cannot turn on and off instantly.

In spectral analysis, windows are used to control 'spectral leakage', which occurs because of the finite time window used for frequency analysis. Spectral leakage generates sidelobes from strong tones that mask low amplitude tones at different frequencies. In a matched filter using Doppler resolving waveforms, reverberation will be much stronger near zero Doppler than at other Doppler frequencies. The waveform windows help keep Doppler sidelobes of reverberation from masking the lower amplitude target echoes that may occur at high Doppler.

Using matched filters based on the narrowband approximation to process echoes with Doppler beyond the narrowband approximation limits will result in correlation loss. The correlation loss will result in a loss of Signal to Noise ratio for these echoes.

This presents a fundamental design decision for a sonar system that needs to process echoes with Doppler on the order of 15 knots or greater. If one uses “narrowband” processing for efficiency, then one has to limit the waveforms to those that satisfy the narrowband approximation. However, as shown in the earlier sections, having a larger bandwidth will reduce the autocorrelation time of the waveform and thus reduce the response to reverberation. This reverberation versus bandwidth property advocates the use of wideband waveforms and hence, broadband matched filtering. There are, however, waveforms that have low correlation loss across all Doppler shifts. These are known as hyperbolic frequency modulation (HFM) waveforms.

So, one can use narrowband processing, and restrict the waveforms to low bandwidth (1 Hertz say) waveforms such as a pulsed sine wave, and wideband waveforms with high Doppler tolerance, such as HFM. If one wants to use waveforms that have Doppler resolving power and high bandwidth, one needs to use broadband matched filtering.

Doppler Sensitive Waveform Matched Filtering

In some cases, one uses a signal that is Doppler sensitive, e.g. the signal ambiguity function $\chi(\tau,\eta)$is a strong function of the Doppler variable. Examples of these waveforms are CW, SFM and comb waveforms. Other waveforms are less sensitive to Doppler effects, beyond a time delay/Doppler coupling effect.

In the cases where one is using Doppler sensitive waveforms, the matched filter is generalized to a matched filter bank, indexed by both time (range) and Doppler (η)

$$m(t,\eta) = \sqrt{\eta} \int_{t}^{t+T/\eta} y(\sigma) r(\eta(\sigma - t)) d\sigma$$

This allows one to search for targets that have relative motion to the source and receivers of the active sonar. When the received signal is a Doppler scaled echo, then the filter that matches the echo Doppler will be a matched filter for that echo, and obey the same SNR properties for echoes embedded in noise as the earlier discussion for stationary targets echoes. The replica is matched to the echo compression and time delay. To ensure energy consistency, we scale the zero Doppler replica $r(t)$ by $\sqrt{\eta}$ when using it for other Doppler hypotheses. This comes from the fact that

$$\int_{0}^{T/\eta} r^2(\eta t)\mathrm{dt} = 1/\eta$$

Now, using the reverberation model as before, we have the following expression for the matched filter bank response to reverberation:

$$m_R(t,\eta) = \sqrt{\eta E_T} \int_{0}^{\infty} A(u) \Gamma(u) \int_{t}^{t+T/\eta} r(\sigma - \tau(u)) r(\eta(\sigma - t)) d\sigma du$$

The integral including the replicas, using the change of variables $\sigma^{'} = \sigma - t$, can be written as

$I_{\delta,\eta} = \int_{0}^{T/\eta} r(\sigma^{'} + \delta) r(\eta\sigma^{'}) d\sigma^{'}$, where $\delta = t - \tau(u)$

We can extend the limits of the integral, because $r(\eta\sigma^{'})$is zero outside the integration limits. Extending limits and changing variables to $\sigma^{''} = \sigma^{'} + \delta$yields

$$I_{\delta,\eta} = \int_{-\infty}^{\infty} r(\sigma'')r(\eta(\sigma'' - \delta))d\sigma'' = \chi(\delta,\eta)/\eta$$

Hence the range Doppler matched filter bank response to reverberation becomes

$$m_R(t,\eta) = \sqrt{E_T}\int_0^{\infty} A(u)\Gamma(u)\chi(t - \tau(u),\eta)du$$

This expression shows that if the target has Doppler, and one uses a replica matched to the target Doppler, further suppression of reverberation is possible. This suppression of reverberation is without loss to matching the target echo or loss in noise limited performance.

Choosing waveforms that are broadband, and with roll-off with respect to Doppler in its signal ambiguity function, will optimize the active sonar's processing in reverberation. For some waveforms (such as the Sinusoidal Frequency Modulation pulse) the signal ambiguity function will have multiple time delay peaks, termed 'fingers', due to the periodic structure of the pulse. Each autocorrelation finger has a time width approximately equal to the signal's bandwidth. This will increase the reverberation response, relative to a waveform with a single autocorrelation response. However, the SFM waveform has a roll-off in Doppler as well. So for targets that have Doppler, there can be a Doppler shift where the roll-off in Doppler more than compensates for the additional autocorrelation peaks.

We will make the assumption that A(u) is wide sense stationary, that is its statistics are invariant over the range of u:

$$E\{A(u)A(v)\} = R_A(u - v)$$

Furthermore, we will assume that A(u) is spatially white, e.g. the scattering elements are uncorrelated with each other:

$$E\{A(u)\,A(v)\} = R_A(u - v) = R_A\delta(u - v)$$

Now, these two assumptions, that the reflection coefficient statistics are independent of range and each differential patch is statistically independent of each other is only an approximation to the real situation. However, these approximations allow one to see the interaction of reverberation and waveform selection.

$$E\{m_R^2(t,\eta)\} = E\left\{E_T\int_0^{\infty} A(u)\Gamma(u)\chi(t - \tau(u),\eta)du\int_0^{\infty} A(\varphi)\Gamma(\varphi)\chi(t - \tau(\varphi),\eta)d\varphi\right\}$$

Rearranging,

$$E\{m_R^2(t,\eta)\} = E_T\int_0^{\infty}\int_0^{\infty} E\{A(u)A(\varphi)\}\Gamma(u)\Gamma(\varphi)\chi(t - \tau(u),\eta)\chi(t - \tau(\varphi),\eta)d\mathrm{u}d\varphi$$

Using the covariance of the scattering elements we get,

$$E\{m_R^2(t,\eta)\} = E_T\int_0^{\infty}\int_0^{\infty} R_A\delta(u - \varphi)\Gamma(u)\Gamma(\varphi)\chi(t - \tau(u),\eta)\chi(t - \tau(\varphi),\eta)d\mathrm{u}d\varphi$$

Or,

$$E\{m_R^2(t,\eta)\} = E_T R_A\int_0^{\infty} |\,\Gamma(u)\,|^2\chi^2(t - \tau(u),\eta)du$$

To see this more clearly, assume that the transmission loss term $\Gamma(u)$is approximately constant over the transmitted signal's correlation time and receiver's beam pattern. Then we obtain

$$E\{m_R^2(t,\eta)\} = E_T R_A |\Gamma(u_0)|^2 \int_0^\infty \chi^2(t-\tau(u),\eta)du$$

Where $u_0$is defined by $\tau(u_0) = t$.

If we assume that the time delay varies smoothly with respect to range, we can replace the integration over u with an integration over time delay τ, where we assume that the chance of variable from u to $\tau$is approximately given by $\tau = 2\mathrm{u}/c$, where c is the speed of sound. This is assuming an approximate monostatic geometry, or that the patch of reverberation is far away relative to the source receiver separation.

We then get

$$E\{m_R^2(t,\eta)\} = E_T R_A |\Gamma(u_0)|^2 c/2 \int_0^\infty \chi^2(t-\tau,\eta)d\tau$$

If we assume that the matched filter time t is greater than the signal duration T, then letting $\tau' = t - \tau$, we obtain

$$E\{m_R^2(t,\eta)\} = E_T R_A |\Gamma(u_0)|^2 c/2 \int_{-\infty}^\infty \chi^2(t-\tau,\eta)d\tau$$

We define the Q-function of the waveform as

$$Q(\eta) = \int_{-\infty}^\infty |\chi(\tau',\eta)|^2 d\tau'$$

Note that $Q(\eta)$has units of seconds^2. We call a waveform with a sharp peak in $Q(\eta)$as a Doppler Sensitive Waveform (DSW). A sine wave pulse will have a sharp peak in $Q(\eta)$ for instance.

When the narrowband ambiguity function is used the Q function is normalized:

$$\int_{-\infty}^\infty Q_{\mathrm{NB}}(\varphi)d\varphi = \int_{-\infty}^\infty |\chi_{\mathrm{NB}}(\tau',\varphi)|^2 d\tau' d\varphi = 1$$

The wideband waveform Q function is approximately normalized to unity.

The reverberation response can be written as

$$E\{m_R^2(t,\eta)\} = E_T R_A |\Gamma(u_0)|^2 Q(\eta) c/2$$

Clearly, the best waveform to use for detection depends on the assumed target velocity. Waveforms such as HFM and LFM have low Q-functions that are relatively constant across Doppler. Doppler sensitive waveforms often have lower Q-functions at higher Doppler shifts than LFM and HFM, much higher Q functions near zero Doppler. To best search for targets, one needs waveforms optimized for both low and high Doppler targets.

So far, this has been a deterministic description of the matched filter response to reverberation.

Channel Doppler Effects on Reverberation

In reality, the reflection coefficient or the transmission loss term will be time varying (as well as spatially varying) because of the surface of the ocean having waves, and the internal thermal structure of the ocean channel will be time varying.

For bottom reverberation, we will assume that the reflection coefficient is time invariant. In shallow water at low frequencies (< 2000 Hz, say) the bottom reverberation dominates over surface reverberation. However, the acoustic propagation through the sound channel and specular reflection from the ocean surface introduces a time varying component to the reverberation formation process.

To derive the results needed for channel Doppler effects, we will restrict ourselves to the narrowband model.

The matched filter is given by:

$$m(t,\varphi) = \int_{t}^{t+T} y(\sigma) r(\sigma - t) e^{-j2\pi\varphi\sigma} d\sigma$$

Since $r(t) = 0$for $t < 0$and $t > T$, we extend the limits of integration for the matched filter response to:

$$m(t,\varphi) = \int_{-\infty}^{\infty} y(\sigma) r(\sigma - t) e^{-j2\pi\varphi\sigma} d\sigma$$

We define the effects of reverberation, targets, clutter and the acoustic channel, via a spreading function $S(\tau,\varphi)$ acting on the transmitted waveform:

$$y(t) = \sqrt{E_T} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} S(\tau,\varphi) e^{j2\pi\varphi t} r(t - \tau) d\tau d\varphi$$

This expression does not include contributions of ambient noise, only scattering phenomena. The spreading function $S(\tau,\varphi)$defines the acoustic scattering, as a function of delay $\tau$and Doppler shift φ for the sonar reception. The spreading function is a random variable, changing due to surface waves and time varying refraction effects (internal waves) in the sound channel.

Target echoes will have a small τregion of non-zero spreading function, $S_{\text{Target}}(\tau,\varphi)$. Reverberation will have an extended τ region with significant $S_{\text{Reverb}}(\tau,\varphi)$. The Doppler shift for both reverberation and targets will be related to receiver and source motion, as well as Doppler spreading due to surface and internal waves. The target will have additional Doppler contributions from its own motion.

Substituting the spreading function description to the sonar response into the matched filter we obtain

$$m(t,\varphi) = \sqrt{E_T} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} S(\tau,\delta) e^{j2\pi\delta\sigma} r(\sigma - \tau) r(\sigma - t) e^{-j2\pi\varphi\sigma} d\sigma d\tau d\delta$$

Which equals

$$m(t,\varphi) = \sqrt{E_T} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} S(\tau,\delta) e^{-j2\pi(\varphi-\delta)\sigma} r(\sigma - \tau) r(\sigma - t) d\sigma d\tau d\delta$$

Letting $\sigma^{'} = \sigma - t$, we obtain

$$m(t,\varphi) = \sqrt{E_T} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} S(\tau,\delta) e^{-j2\pi(\varphi-\delta)(\sigma'+\tau)} r(\sigma') r(\sigma' - (t-\tau)) d\sigma d\tau d\delta$$

Using the definition of the narrowband ambiguity function, the matched filter response becomes

$$m(t,\varphi) = \sqrt{E_T} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} S(\tau,\delta) e^{-j2\pi(\varphi-\delta)\tau} \chi(t-\tau,\varphi-\delta) d\tau d\delta$$

The response of the matched filter is a “twisted convolution” of the spreading function and the waveform ambiguity function. The exponential $e^{-j2\pi(\varphi-\delta)}$performs the twisting. Note that if the waveform ambiguity function was “perfect”, that is a single peak,

$$\chi(\tau,\varphi) = \delta(\tau)\delta(\varphi)$$

Then the matched filter response would become:

$$m(t,\varphi) = \sqrt{E_T} S(t,\varphi) + n(t,\varphi)$$

Where $n(t,\varphi)$ is the response of the matched filter to ambient noise. In this sense, the matched filter is estimating the spreading function of the channel, with targets, clutter and reverberation all part of the spreading function. Note, however that $\chi(0,0) = 1$, so the ambiguity function cannot become a delta function.

Now, the power output of the matched filter is desired, so that Signal to Interference Ratios and similar quantities can be predicted. We will make statistical assumptions about the spreading function. The assumptions are that the spreading function is wide sense stationary and uncorrelated. This implies that the signals being processed are statistically stationary and that the scatterers are uncorrelated; so that (Van Trees, III, Ch 13):

$$E\{S(\tau,\varphi) S(\tau',\varphi')\} = R_{\text{SS}}(\tau,\varphi)\delta(\tau-\tau')\delta(\varphi-\varphi')$$

$R_{\text{SS}}(\tau,\varphi)$is known as the scattering function of the active sonar scenario. The description of the target, reverberation and clutter statistics are captured in this expression.

Using this definition, we obtain for the power of the matched filter:

$$E\{m(t,\varphi) m(t,\varphi)\} =$$

$$E_T \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} S(\tau,\delta) e^{-j2\pi(\varphi-\delta)\tau} \chi(t-\tau,\varphi-\delta) d\tau d\delta \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} S(\tau',\delta') e^{j2\pi(\varphi-\delta')\tau'} \chi(t-\tau',\varphi-\delta') d\tau' d\delta'$$

Which becomes

$$E\left\{ |\, m(t,\varphi)\, |^2 \right\} = E_T \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} R_{\text{SS}}(\tau,\delta) |\, \chi(t-\tau,\varphi-\delta)\, |^2 d\tau d\delta$$

Using two dimensional convolution notation **, this expression for the matched filter power $P(t,\varphi)$ becomes

$$P(t,\varphi) = E\left\{ |\, m(t,\varphi)\, |^2 \right\} = E_T R_{\text{SS}}(t,\varphi) ** |\, \chi(t,\varphi)\, |^2$$

Now, let us model the acoustic sonar problem as target, clutter/reverberation and noise. The matched filter power response to ambient noise was shown earlier to be AN_0. The overall signal to interference ratio, for a target at range $t$and Doppler φ is

$$\mathrm{SIR} = \frac{P_{\mathrm{Target}}(t,\varphi)}{P_{\mathrm{Reverb/clutter}}(t,\varphi)+\mathrm{AN}_0}$$

Which becomes,

$$\mathrm{SIR}(t,\varphi) = \frac{R_{\mathrm{SS}}^{\mathrm{Target}}(t,\varphi)^{**}|\chi(t,\varphi)|^2}{R_{\mathrm{SS}}^{\mathrm{Reverb/Clutter}}(t,\varphi)^{**}|\chi(t,\varphi)|^2+\frac{\mathrm{AN}_0}{E_T}}$$

This expression can be simplified for different target and scattering conditions. For a point target at range $t_0$and Doppler φ_0, the target scattering function becomes

$$R_{\mathrm{SS}}^{\mathrm{Target}}(t,\varphi) = S\delta(t-t_0)\delta(\varphi-\varphi_0).$$

Often, we can assume that the reverberation scattering function is constant in the vicinity of the target, so that

$$R_{\mathrm{SS}}^{\mathrm{Reverb/Clutter}}(t,\varphi) = R_{t_0}Q_{\mathrm{Revrb}}(\varphi)$$

$Q_{\mathrm{Revrb}}(\varphi)$ describes the Doppler roll-off of the reverberation. It will be affected by the source, receiver and ocean motion. In this characterization, we are ignoring “discrete clutter”, e.g. target like responses from bottom features.

We will assume that $Q_{\mathrm{Revrb}}(\varphi)$is normalized, so that:

$$\int Q_{\mathrm{Revrb}}(\varphi)d\varphi = 1$$

With these assumptions we obtain

$$\mathrm{SIR}(t,\varphi) = \frac{S^{\mathrm{Target}}(t_0,\varphi_0)|\chi(t-t_0,\varphi-\varphi_0)|^2}{R_{t_0}\iint Q_{\mathrm{Reverb}}(\varphi-\delta)|\chi(\tau,\varphi)|^2 d\tau d\delta+\frac{\mathrm{AN}_0}{E_T}}$$

Note that when the matched filter is matched in time and Doppler, then $t = t_0$and $\varphi = \varphi_0$, and the numerator is maximized:

$$\mathrm{SIR}(t_0,\varphi_0) = \frac{S^{\mathrm{Target}}(t_0,\varphi_0)}{R_{t_0}\iint Q_{\mathrm{Reverb}}(\varphi_0-\delta)|\chi(\tau,\varphi_0)|^2 d\tau d\delta+\frac{\mathrm{AN}_0}{E_T}}$$

The denominator can be simplified further by using the definition of the waveform Q function:

$$Q(\varphi) = \int |\,\chi(\tau,\varphi)\,|^2 d\tau$$

Using this definition, we obtain:

$$\mathrm{SIR}(t,\varphi) = \frac{S^{\mathrm{Target}}(t_0,\varphi_0)|\chi(t-t_0,\varphi-\varphi_0)|^2}{R_{t_0}Q_{\mathrm{Reverb}}(\varphi)*Q(\varphi)+\frac{\mathrm{AN}_0}{E_T}}$$

When using a waveform with a Q function much wider than the reverberation Q function, (such as an HFM), the waveform Q function can be replaced by a constant, such as Q_{HFM}. The convolution with the

reverberation Q function becomes the constant Q_{HFM}(because of the normalization of Q_{Reverb}) :

$$Q_{\mathrm{Reverb}}(\varphi) * Q(\varphi) = Q_{\mathrm{HFM}} \int Q_{\mathrm{Reverb}}(\varphi - \varphi^{'}) d\varphi^{'} = Q_{\mathrm{HFM}}$$

With this approximation the signal to interference ratio becomes:

$$\mathrm{SIR}_{\mathrm{HFM}}(t,\varphi) = \frac{S^{\mathrm{Target}}(t_0,\varphi_0)|\chi(t-t_0,\varphi-\varphi_0)|^2}{R_{t_0} Q_{\mathrm{HFM}} + \frac{\mathrm{AN}_0}{E_T}}$$

When the waveform is Doppler sensitive, and it's Q function is narrower than the Reverberation Q function, we can approximate the waveform Q function by a rectangle of height T and width 1/T centered at zero Doppler. Than the convolution becomes:

$$Q_{\mathrm{Reverb}}(\varphi) * Q(\varphi) = \int Q_{\mathrm{Reverb}}(\varphi - \varphi^{'}) Q_{\mathrm{DSW}}(\varphi^{'}) d\varphi^{'} \approx \int_{-1/2\mathrm{T}}^{1/2\mathrm{T}} Q_{\mathrm{Reverb}}(\varphi - \varphi^{'}) \mathrm{T} \mathrm{d}\varphi^{'} \approx Q_{\mathrm{Reverb}}(\varphi)$$

Therefore, the signal to interference ratio becomes

$$\mathrm{SIR}_{\mathrm{DSW}}(t,\varphi) = \frac{S^{\mathrm{Target}}(t_0,\varphi_0)|\chi(t-t_0,\varphi-\varphi_0)|^2}{R_{t_0} Q_{\mathrm{Reverb}}(\varphi) + \frac{\mathrm{AN}_0}{E_T}}$$

We see that the best waveform for enhancing signal to interference ratio depends on the environmental Q function, and the assumed target Doppler.

The Active Sonar Equation

The active sonar equation expresses the signal excess (SE) which is the part of the target signal to noise ratio that exceeds the sonar's detection threshold (DT). In decibel quantities, it is given by:

$$\mathrm{SE} = \mathrm{ESL} + \mathrm{TS} - \mathrm{TL}_{\mathrm{ST}} - \mathrm{TL}_{\mathrm{TR}} - (\mathrm{RL}_0 \oplus \mathrm{AN}_0) - \mathrm{DT}$$

We are assuming that the active sonar uses a matched filter for detection. In the sonar equation, the transmitted energy signal level (ESL) is the sound pressure squared and integrated over the transmitted pulse length. The energy of the received echo (known as the echo energy level) is $10*\log10(E_R) = \mathrm{ESL} + \mathrm{TS} - \mathrm{TL}_{\mathrm{ST}} - \mathrm{TL}_{\mathrm{TR}}$. Note that the echo energy level can be computed using a received echo pulse length, which due to sound channel dispersion can be longer than the transmitted pulse.

Using the properties of matched filters, the matched filter generates an output due to the target echo with peak power level of $\mathrm{ESL} + \mathrm{TS} - \mathrm{TL}_{\mathrm{ST}} - \mathrm{TL}_{\mathrm{TR}}$.

$\mathrm{AN}_0$is the ambient noise level in a 1 Hz band, and is assumed to be constant across the matched filter's bandwidth, e.g. it is twice the spectral density of the ambient noise. AN_0 and spectral density have units of $\mathrm{Pa}^2/\mathrm{Hz}$, which is an energy quantity.

The matched filter noise output is $\mathrm{RL}_0 \oplus \mathrm{AN}_0$. The operation $\oplus$ corresponds to power addition. Power addition converts the quantities back to units of power (Pa^2, volts^2, etc), adds the two power like quantities, and then reconverts back into decibel.

$$\mathrm{RL}_0 \oplus \mathrm{AN}_0 = 10\log_{10}\, 10^{\mathrm{RL}_0/10} + 10^{\mathrm{AN}_0/10}$$

$\mathrm{RL}_0$is the reverberation level in a 1 Hz band, or equivalently, the reverberation level when measured at the output of the matched filter.